

VISIT...

LANZAROTE
Caliente.COM

The Transhumanist FAQ

– A General Introduction –

Version 2.1 (2003)

Nick Bostrom*
Faculty of Philosophy
Oxford University

* Please see endnote for document history and acknowledgments.

Published by the *World Transhumanist Association*



www.nickbostrom.com

CONTENTS

1	GENERAL QUESTIONS ABOUT TRANSHUMANISM.....	4
1.1	What is transhumanism?	4
1.2	What is a posthuman?	5
1.3	What is a transhuman?	6
2	TECHNOLOGIES AND PROJECTIONS	7
2.1	Biotechnology, genetic engineering, stem cells, and cloning – what are they and what are they good for?	7
2.2	What is molecular nanotechnology?	9
2.3	What is superintelligence?	12
2.4	What is virtual reality?	14
2.5	What is cryonics? Isn't the probability of success too small?.....	15
2.6	What is uploading?.....	17
2.7	What is the singularity?.....	19
3	SOCIETY AND POLITICS.....	20
3.1	Will new technologies only benefit the rich and powerful?.....	20
3.2	Do transhumanists advocate eugenics?	21
3.3	Aren't these future technologies very risky? Could they even cause our extinction?	22
3.4	If these technologies are so dangerous, should they be banned? What can be done to reduce the risks?	25
3.5	Shouldn't we concentrate on current problems such as improving the situation of the poor, rather than putting our efforts into planning for the “far” future? 27	
3.6	Will extended life worsen overpopulation problems?.....	28
3.7	Is there any ethical standard by which transhumanists judge “improvement of the human condition”?	31
3.8	What kind of society would posthumans live in?	32
3.9	Will posthumans or superintelligent machines pose a threat to humans who aren't augmented?	33
4	TRANSHUMANISM AND NATURE.....	34
4.1	Why do transhumanists want to live longer?	34
4.2	Isn't this tampering with nature?.....	35
4.3	Will transhuman technologies make us inhuman?.....	36
4.4	Isn't death part of the natural order of things?	36
4.5	Are transhumanist technologies environmentally sound?.....	38
5	TRANSHUMANISM AS A PHILOSOPHICAL AND CULTURAL VIEWPOINT.....	38
5.1	What are the philosophical and cultural antecedents of transhumanism?.....	38
5.2	What currents are there within transhumanism? Is extropianism the same as transhumanism?.....	44
5.3	How does transhumanism relate to religion?	45
5.4	Won't things like uploading, cryonics, and AI fail because they can't preserve or create the soul?	46
5.5	What kind of transhumanist art is there?.....	47

6	PRACTICALITIES	48
6.1	What are the reasons to expect all these changes?.....	48
6.2	Won't these developments take thousands or millions of years?	49
6.3	What if it doesn't work?.....	50
6.4	How can I use transhumanism in my own life?	51
6.5	How could I become a posthuman?	51
6.6	Won't it be boring to live forever in a perfect world?	52
6.7	How can I get involved and contribute?.....	53
7	ACKNOWLEDGEMENTS AND DOCUMENT HISTORY.....	54

1 GENERAL QUESTIONS ABOUT TRANSHUMANISM

1.1 What is transhumanism?

Transhumanism is a way of thinking about the future that is based on the premise that the human species in its current form does not represent the end of our development but rather a comparatively early phase. We formally define it as follows:

- (1) The intellectual and cultural movement that affirms the possibility and desirability of fundamentally improving the human condition through applied reason, especially by developing and making widely available technologies to eliminate aging and to greatly enhance human intellectual, physical, and psychological capacities.
- (2) The study of the ramifications, promises, and potential dangers of technologies that will enable us to overcome fundamental human limitations, and the related study of the ethical matters involved in developing and using such technologies.

Transhumanism can be viewed as an extension of humanism, from which it is partially derived. Humanists believe that humans matter, that individuals matter. We might not be perfect, but we can make things better by promoting rational thinking, freedom, tolerance, democracy, and concern for our fellow human beings. Transhumanists agree with this but also emphasize what we have the potential to become. Just as we use rational means to improve the human condition and the external world, we can also use such means to improve ourselves, the human organism. In doing so, we are not limited to traditional humanistic methods, such as education and cultural development. We can also use technological means that will eventually enable us to move beyond what some would think of as “human”.

It is not our human shape or the details of our current human biology that define what is valuable about us, but rather our aspirations and ideals, our experiences, and the kinds of lives we lead. To a transhumanist, progress occurs when more people become more able to shape themselves, their lives, and the ways they relate to others, in accordance with their own deepest values. Transhumanists place a high value on autonomy: the ability and right of individuals to plan and choose their own lives. Some people may of course, for any number of reasons, choose to forgo the opportunity to use technology to improve themselves. Transhumanists seek to create a world in which autonomous individuals may choose to remain unenhanced or choose to be enhanced and in which these choices will be respected.

Through the accelerating pace of technological development and scientific understanding, we are entering a whole new stage in the history of the human species. In the relatively

near future, we may face the prospect of real artificial intelligence. New kinds of cognitive tools will be built that combine artificial intelligence with interface technology. Molecular nanotechnology has the potential to manufacture abundant resources for everybody and to give us control over the biochemical processes in our bodies, enabling us to eliminate disease and unwanted aging. Technologies such as brain-computer interfaces and neuropharmacology could amplify human intelligence, increase emotional well-being, improve our capacity for steady commitment to life projects or a loved one, and even multiply the range and richness of possible emotions. On the dark side of the spectrum, transhumanists recognize that some of these coming technologies could potentially cause great harm to human life; even the survival of our species could be at risk. Seeking to understand the dangers and working to prevent disasters is an essential part of the transhumanist agenda.

Transhumanism is entering the mainstream culture today, as increasing numbers of scientists, scientifically literate philosophers, and social thinkers are beginning to take seriously the range of possibilities that transhumanism encompasses. A rapidly expanding family of transhumanist groups, differing somewhat in flavor and focus, and a plethora of discussion groups in many countries around the world, are gathered under the umbrella of the World Transhumanist Association, a non-profit democratic membership organization.

References:

World Transhumanist Association. <http://www.transhumanism.org>

1.2 What is a posthuman?

It is sometimes useful to talk about possible future beings whose basic capacities so radically exceed those of present humans as to be no longer unambiguously human by our current standards. The standard word for such beings is “posthuman”. (Care must be taken to avoid misinterpretation. “Posthuman” does not denote just anything that happens to come after the human era, nor does it have anything to do with the “posthumous”. In particular, it does *not* imply that there are no humans anymore.)

Many transhumanists wish to follow life paths which would, sooner or later, require growing into posthuman persons: they yearn to reach intellectual heights as far above any current human genius as humans are above other primates; to be resistant to disease and impervious to aging; to have unlimited youth and vigor; to exercise control over their own desires, moods, and mental states; to be able to avoid feeling tired, hateful, or irritated about petty things; to have an increased capacity for pleasure, love, artistic appreciation, and serenity; to experience novel states of consciousness that current human brains cannot access. It seems likely that the simple fact of living an indefinitely long, healthy, active life would take anyone to posthumanity if they went on accumulating memories, skills, and intelligence.

Posthumans could be completely synthetic artificial intelligences, or they could be enhanced uploads [see “What is uploading?”], or they could be the result of making many smaller but cumulatively profound augmentations to a biological human. The latter

alternative would probably require either the redesign of the human organism using advanced nanotechnology or its radical enhancement using some combination of technologies such as genetic engineering, psychopharmacology, anti-aging therapies, neural interfaces, advanced information management tools, memory enhancing drugs, wearable computers, and cognitive techniques.

Some authors write as though simply by changing our self-conception, we have become or could become posthuman. This is a confusion or corruption of the original meaning of the term. The changes required to make us posthuman are too profound to be achievable by merely altering some aspect of psychological theory or the way we think about ourselves. Radical technological modifications to our brains and bodies are needed.

It is difficult for us to imagine what it would be like to be a posthuman person. Posthumans may have experiences and concerns that we cannot fathom, thoughts that cannot fit into the three-pound lumps of neural tissue that we use for thinking. Some posthumans may find it advantageous to jettison their bodies altogether and live as information patterns on vast super-fast computer networks. Their minds may be not only more powerful than ours but may also employ different cognitive architectures or include new sensory modalities that enable greater participation in their virtual reality settings. Posthuman minds might be able to share memories and experiences directly, greatly increasing the efficiency, quality, and modes in which posthumans could communicate with each other. The boundaries between posthuman minds may not be as sharply defined as those between humans.

Posthumans might shape themselves and their environment in so many new and profound ways that speculations about the detailed features of posthumans and the posthuman world are likely to fail.

1.3 What is a transhuman?

In its contemporary usage, “transhuman” refers to an intermediary form between the human and the posthuman [see “What is a posthuman?”]. One might ask, given that our current use of e.g. medicine and information technology enable us to routinely do many things that would have astonished humans living in ancient times, whether we are not already transhuman? The question is a provocative one, but ultimately not very meaningful; the concept of the transhuman is too vague for there to be a definite answer.

A *transhumanist* is simply someone who advocates transhumanism [see “What is transhumanism?”]. It is a common error for reporters and other writers to say that transhumanists “claim to be transhuman” or “call themselves transhuman”. To adopt a philosophy which says that someday everyone ought to have the chance to grow beyond present human limits is clearly not to say that one is better or somehow currently “more advanced” than one’s fellow humans.

The etymology of the term “transhuman” goes back to the futurist FM-2030 (also known as F. M. Estfandiary), who introduced it as shorthand for “transitional human”. Calling transhumans the “earliest manifestation of new evolutionary beings,” FM maintained that signs of transhumanity included prostheses, plastic surgery, intensive use of telecommunications, a cosmopolitan outlook and a globetrotting lifestyle, androgyny, mediated reproduction (such as *in vitro* fertilization), absence of religious beliefs, and a rejection of traditional family values. However, FM’s diagnostics are of dubious validity. It is unclear why anybody who has a lot of plastic surgery or a nomadic lifestyle is any closer to becoming a posthuman than the rest of us; nor, of course, are such persons necessarily more admirable or morally commendable than others. In fact, it is perfectly possible to be a transhuman – or, for that matter, a transhumanist – and still embrace most traditional values and principles of personal conduct.

References:

FM-2030. *Are You a Transhuman?* (New York: Warner Books, 1989).

2 TECHNOLOGIES AND PROJECTIONS

2.1 Biotechnology, genetic engineering, stem cells, and cloning – what are they and what are they good for?

Biotechnology is the application of techniques and methods based on the biological sciences. It encompasses such diverse enterprises as brewing, manufacture of human insulin, interferon, and human growth hormone, medical diagnostics, cell cloning and reproductive cloning, the genetic modification of crops, bioconversion of organic waste and the use of genetically altered bacteria in the cleanup of oil spills, stem cell research, and much more. Genetic engineering is the area of biotechnology concerned with the directed alteration of genetic material.

Biotechnology already has countless applications in industry, agriculture, and medicine. It is a hotbed of research. The completion of the human genome project – a “rough draft” of the entire human genome was published in the year 2000 – was a scientific milestone by anyone’s standards. Research is now shifting to decoding the functions and interactions of all these different genes and to developing applications based on this information.

The potential medical benefits are too many to list; researchers are working on every common disease, with varying degrees of success. Progress takes place not only in the development of drugs and diagnostics but also in the creation of better tools and research methodologies, which in turn accelerates progress. When considering what developments are likely over the long term, such improvements in the research process itself must be factored in. The human genome project was completed ahead of schedule, largely because the initial predictions underestimated the degree to which instrumentation

technology would improve during the course of the project. At the same time, one needs to guard against the tendency to hype every latest advance. (Remember all those breakthrough cancer cures that we never heard of again?) Moreover, even in cases where the early promise is borne out, it usually takes ten years to get from proof-of-concept to successful commercialization.

Genetic therapies are of two sorts: somatic and germ-line. In somatic gene therapy, a virus is typically used as a vector to insert genetic material into the cells of the recipient's body. The effects of such interventions do not carry over into the next generation. Germ-line genetic therapy is performed on sperm or egg cells, or on the early zygote, and can be inheritable. (Embryo screening, in which embryos are tested for genetic defects or other traits and then selectively implanted, can also count as a kind of germ-line intervention.) Human gene therapy, except for some forms of embryo screening, is still experimental. Nonetheless, it holds promise for the prevention and treatment of many diseases, as well as for uses in enhancement medicine. The potential scope of genetic medicine is vast: virtually all disease and all human traits – intelligence, extroversion, conscientiousness, physical appearance, etc. – involve genetic predispositions. Single-gene disorders, such as cystic fibrosis, sickle cell anemia, and Huntington's disease are likely to be among the first targets for genetic intervention. Polygenic traits and disorders, ones in which more than one gene is implicated, may follow later (although even polygenic conditions can sometimes be influenced in a beneficial direction by targeting a single gene).

Stem cell research, another scientific frontier, offers great hopes for regenerative medicine. Stem cells are undifferentiated (unspecialized) cells that can renew themselves and give rise to one or more specialized cell types with specific functions in the body. By growing such cells in culture, or steering their activity in the body, it will be possible to grow replacement tissues for the treatment of degenerative disorders, including heart disease, Parkinson's, Alzheimer's, diabetes, and many others. It may also be possible to grow entire organs from stem cells for use in transplantation. Embryonic stem cells seem to be especially versatile and useful, but research is also ongoing into adult stem cells and the "reprogramming" of ordinary cells so that they can be turned back into stem cells with pluripotent capabilities.

The term "human cloning" covers both therapeutic and reproductive uses. In therapeutic cloning, a preimplantation embryo (also known as a "blastocyst" – a hollow ball consisting of 30-150 undifferentiated cells) is created via cloning, from which embryonic stem cells could be extracted and used for therapy. Because these cloned stem cells are genetically identical to the patient, the tissues or organs they would produce could be implanted without eliciting an immune response from the patient's body, thereby overcoming a major hurdle in transplant medicine. Reproductive cloning, by contrast, would mean the birth of a child who is genetically identical to the cloned parent: in effect, a younger identical twin.

Everybody recognizes the benefit to ailing patients and their families that come from curing specific diseases. Transhumanists emphasize that, in order to seriously prolong the

healthy life span, we also need to develop ways to slow aging or to replace senescent cells and tissues. Gene therapy, stem cell research, therapeutic cloning, and other areas of medicine that have the potential to deliver these benefits deserve a high priority in the allocation of research monies.

Biotechnology can be seen as a special case of the more general capabilities that nanotechnology will eventually provide [see “What is molecular nanotechnology?”].

2.2 What is molecular nanotechnology?

Molecular nanotechnology is an anticipated manufacturing technology that will make it possible to build complex three-dimensional structures to atomic specification using chemical reactions directed by nonbiological machinery. In molecular manufacturing, each atom would go to a selected place, bonding with other atoms in a precisely designated manner. Nanotechnology promises to give us thorough control of the structure of matter.

Since most of the stuff around us and inside us is composed of atoms and gets its characteristic properties from the placement of these atoms, the ability to control the structure of matter on the atomic scale has many applications. As K. Eric Drexler wrote in *Engines of Creation*, the first book on nanotechnology (published in 1986):

Coal and diamonds, sand and computer chips, cancer and healthy tissue: throughout history, variations in the arrangement of atoms have distinguished the cheap from the cherished, the diseased from the healthy. Arranged one way, atoms make up soil, air, and water arranged another, they make up ripe strawberries. Arranged one way, they make up homes and fresh air; arranged another, they make up ash and smoke.

Nanotechnology, by making it possible to rearrange atoms effectively, will enable us to transform coal into diamonds, sand into supercomputers, and to remove pollution from the air and tumors from healthy tissue.

Central to Drexler’s vision of nanotechnology is the concept of the *assembler*. An assembler would be a molecular construction device. It would have one or more submicroscopic robotic arms under computer control. The arms would be capable of holding and placing reactive compounds so as to positionally control the precise location at which a chemical reaction takes place. The assembler arms would grab a molecule (but *not* necessarily individual atoms) and add it to a work-piece, constructing an atomically precise object step by step. An advanced assembler would be able to make almost any chemically stable structure. In particular, it would be able to make a copy of itself. Since assemblers could replicate themselves, they would be easy to produce in large quantities.

There is a biological parallel to the assembler: the ribosome. Ribosomes are the tiny construction machines (a few thousand cubic nanometers big) in our cells that

manufacture all the proteins used in all living things on Earth. They do this by assembling amino acids, one by one, into precisely determined sequences. These structures then fold up to form a protein. The blueprint that specifies the order of amino acids, and thus indirectly the final shape of the protein, is called messenger RNA. The messenger RNA is in turn determined by our DNA, which can be viewed (somewhat simplistically) as an instruction tape for protein synthesis. Nanotechnology will generalize the ability of ribosomes so that virtually any chemically stable structure can be built, including devices and materials that resemble nothing in nature.

Mature nanotechnology will transform manufacturing into a software problem. To build something, all you will need is a detailed design of the object you want to make and a sequence of instructions for its construction. Rare or expensive raw materials are generally unnecessary; the atoms required for the construction of most kinds of nanotech devices exist in abundance in nature. Dirt, for example, is full of useful atoms.

By working in large teams, assemblers and more specialized nanomachines will be able to build large objects quickly. Consequently, while nanomachines may have features on the scale of a billionth of a meter – a nanometer – the products could be as big as space vehicles or even, in a more distant future, the size of planets.

Because assemblers will be able to copy themselves, nanotech products will have low marginal production costs – perhaps on the same order as familiar commodities from nature's own self-reproducing molecular machinery such as firewood, hay, or potatoes. By ensuring that each atom is properly placed, assemblers would manufacture products of high quality and reliability. Leftover molecules would be subject to this strict control, making the manufacturing process extremely clean.

The speed with which designs and instruction lists for making useful objects can be developed will determine the speed of progress after the creation of the first full-blown assembler. Powerful software for molecular modeling and design will accelerate development, possibly assisted by specialized engineering AI. Another accessory that might be especially useful in the early stages after the assembler-breakthrough is the *disassembler*, a device that can disassemble an object while creating a three-dimensional map of its molecular configuration. Working in concert with an assembler, it could function as a kind of 3D Xerox machine: a device for making atomically exact replicas of almost any existing solid object within reach.

Molecular nanotechnology will ultimately make it possible to construct compact computing systems performing at least 10^{21} operations per second; machine parts of any size made of nearly flawless diamond; cell-repair machines that can enter cells and repair most kinds of damage, in all likelihood including frostbite [see "What is cryonics? Isn't the probability of success too small?"]; personal manufacturing and recycling appliances; and automated production systems that can double capital stock in a few hours or less. It is also likely to make uploading possible [see "What is uploading?"].

A key challenge in realizing these prospects is the bootstrap problem: how to build the first assembler. There are several promising routes. One is to improve current proximal probe technology. An atomic force microscope can drag individual atoms along a surface. Two physicists at IBM Almaden Labs in California illustrated this in 1989 when they used such a microscope to arrange 35 xenon atoms to spell out the trademark “I-B-M”, creating the world’s smallest logo. Future proximal probes might have more degrees of freedom and the ability to pick up and deposit reactive compounds in a controlled fashion.

Another route to the first assembler is synthetic chemistry. Cleverly designed chemical building blocks might be made to self-assemble in solution phase into machine parts. Final assembly of these parts might then be made with a proximal probe.

Yet another route is biochemistry. It might be possible to use ribosomes to make assemblers of more generic capabilities. Many biomolecules have properties that might be explored in the early phases of nanotechnology. For example, interesting structures, such as branches, loops, and cubes, have been made by DNA. DNA could also serve as a “tag” on other molecules, causing them to bind only to designated compounds displaying a complementary tag, thus providing a degree of control over what molecular complexes will form in a solution.

Combinations of these approaches are of course also possible. The fact that there are multiple promising routes adds to the likelihood that success will eventually be attained.

That assemblers of general capabilities are consistent with the laws of chemistry was shown by Drexler in his technical book *Nanosystems* in 1992. This book also established some lower bounds on the capabilities of mature nanotechnology. Medical applications of nanotechnology were first explored in detail by Robert A. Freitas Jr. in his monumental work *Nanomedicine*, the first volume of which came out in 1999. Today, nanotech is a hot research field. The U.S. government spent more than 600 million dollars on its National Nanotechnology Initiative in 2002. Other countries have similar programs, and private investment is ample. However, only a small part of the funding goes to projects of direct relevance to the development of assembler-based nanotechnology; most of it is for more humdrum, near-term objectives.

While it seems fairly well established that molecular nanotechnology is in principle possible, it is harder to determine how long it will take to develop. A common guess among the cognoscenti is that the first assembler may be built around the year 2018, give or take a decade, but there is large scope for diverging opinion on the upper side of that estimate.

Because the ramifications of nanotechnology are immense, it is imperative that serious thought be given to this topic now. If nanotechnology were to be abused the consequences could be devastating. Society needs to prepare for the assembler breakthrough and do advance planning to minimize the risks associated with it [see e.g. “Aren’t these future technologies very risky? Could they even cause our extinction?”].

Several organizations are working to preparing the world for nanotechnology, the oldest and largest being the Foresight Institute.

References:

Drexler, E. *The Engines of Creation: The Coming Era of Nanotechnology*. (New York: Anchor Books, 1986). <http://www.foresight.org/EOC/index.html>

Drexler, E. *Nanosystems: Molecular Machinery, Manufacturing, and Computation*. (New York: John Wiley & Sons, Inc., 1992).

Freitas, Jr., R. A. *Nanomedicine, Volume I: Basic Capabilities*. (Georgetown, Texas: Landes Bioscience, 1999).

Foresight Institute. <http://www.foresight.org>

2.3 What is superintelligence?

A superintelligent intellect (a superintelligence, sometimes called “ultraintelligence”) is one that has the capacity to radically outperform the best human brains in practically every field, including scientific creativity, general wisdom, and social skills.

Sometimes a distinction is made between weak and strong superintelligence. *Weak superintelligence* is what you would get if you could run a human intellect at an accelerated clock speed, such as by uploading it to a fast computer [see “What is uploading?”]. If the upload’s clock-rate were a thousand times that of a biological brain, it would perceive reality as being slowed down by a factor of a thousand. It would think a thousand times more thoughts in a given time interval than its biological counterpart.

Strong superintelligence refers to an intellect that is not only faster than a human brain but also smarter in a qualitative sense. No matter how much you speed up your dog’s brain, you’re not going to get the equivalent of a human intellect. Analogously, there might be kinds of smartness that wouldn’t be accessible to even very fast human brains given their current capacities. Something as simple as increasing the size or connectivity of our neuronal networks *might* give us some of these capacities. Other improvements may require wholesale reorganization of our cognitive architecture or the addition of new layers of cognition on top of the old ones.

However, the distinction between weak and strong superintelligence may not be clear-cut. A sufficiently long-lived human who didn’t make any errors and had a sufficient stack of scrap paper at hand could in principle compute any Turing computable function. (According to Church’s thesis, the class of Turing computable functions is identical to the class of physically computable functions.)

Many but not all transhumanists expect that superintelligence will be created within the first half of this century. Superintelligence requires two things: hardware and software.

Chip-manufacturers planning the next generation of microprocessors commonly rely on a well-known empirical regularity known as Moore's Law. In its original 1965-formulation by Intel co-founder Gordon Moore, it stated that the number of components on a chip doubled every year. In contemporary use, the "law" is commonly understood as referring more generally to a doubling of computing power, or of computing power per dollar. For the past couple of years, the doubling time has hovered between 18 months and two years.

The human brain's processing power is difficult to determine precisely, but common estimates range from 10^{14} instructions per second (IPS) up to 10^{17} IPS or more. The lower estimate, derived by Carnegie Mellon robotics professor Hans Moravec, is based on the computing power needed to replicate the signal processing performed by the human retina and assumes a significant degree of software optimization. The 10^{17} IPS estimate is obtained by multiplying the number of neurons in a human brain (~100 billion) with the average number of synapses per neuron (~1,000) and with the average spike rate (~100 Hz), and assuming ~10 instructions to represent the effect on one action potential traversing one synapse. An even higher estimate would be obtained e.g. if one were to suppose that functionally relevant and computationally intensive processing occurs within compartments of a dendrite tree.

Most experts, Moore included, think that computing power will continue to double about every 18 months for at least another two decades. This expectation is based in part on extrapolation from the past and in part on consideration of developments currently underway in laboratories. The fastest computer under construction is IBM's Blue Gene/L, which when it is ready in 2005 is expected to perform $\sim 2 \times 10^{14}$ IPS. Thus it appears quite likely that human-equivalent hardware will have been achieved within not much more than a couple of decades.

How long it will take to solve the software problem is harder to estimate. One possibility is that progress in computational neuroscience will teach us about the computational architecture of the human brain and what learning rules it employs. We can then implement the same algorithms on a computer. In this approach, the superintelligence would not be completely specified by the programmers but would instead have to grow by learning from experience the same way a human infant does. An alternative approach would be to use genetic algorithms and methods from classical AI. This might result in a superintelligence that bears no close resemblance to a human brain. At the opposite extreme, we could seek to create a superintelligence by uploading a human intellect and then accelerating and enhancing it [see "What is uploading?"]. The outcome of this might be a superintelligence that is a radically upgraded version of one particular human mind.

The arrival of superintelligence will clearly deal a heavy blow to anthropocentric worldviews. Much more important than its philosophical implications, however, would be its practical effects. Creating superintelligence may be the last invention that humans will ever need to make, since superintelligences could themselves take care of further scientific and technological development. They would do so more effectively than humans. Biological humanity would no longer be the smartest life form on the block.

The prospect of superintelligence raises many big issues and concerns that we should think deeply about in advance of its actual development. The paramount question is: What can be done to maximize the chances that the arrival of superintelligence will benefit rather than harm us? The range of expertise needed to address this question extends far beyond the community of AI researchers. Neuroscientists, economists, cognitive scientists, computer scientists, philosophers, ethicists, sociologists, science-fiction writers, military strategists, politicians, legislators, and many others will have to pool their insights if we are to deal wisely with what may be the most important task our species will ever have to tackle.

Many transhumanists would like to become superintelligent themselves. This is obviously a long-term and uncertain goal, but it might be achievable either through uploading and subsequent enhancement or through the gradual augmentation of our biological brains, by means of future nootropics (cognitive enhancement drugs), cognitive techniques, IT tools (e.g. wearable computers, smart agents, information filtering systems, visualization software, etc.), neural-computer interfaces, or brain implants.

References:

Moravec, H. *Mind Children* (Harvard: Harvard University Press, 1989).

Bostrom, N. "How Long Before Superintelligence?" *International Journal of Futures Studies*. Vol. 2. (1998).

2.4 What is virtual reality?

A virtual reality is a simulated environment that your senses perceive as real.

Theatre, opera, cinema, television can be regarded as precursors to virtual reality. The degree of immersion (the feeling of "being there") that you experience when watching television is quite limited. Watching football on TV doesn't really compare to being in the stadium. There are several reasons for this. For starters, even a big screen doesn't fill up your entire visual field. The number of pixels even on high-resolution screens is also too small (typically 1280*1224 rather than about 5000*5000 as would be needed in a flawless wide-angle display). Further, 3D vision is lacking, as is position tracking and focus effects (in reality, the picture on your retina changes continually as your head and eyeballs are moving). To achieve greater realism, a system should ideally include more sensory modalities, such as 3D sound (through headphones) to hear the crowd roaring, and tactile stimulation through a whole-body haptic interface so that you don't have to miss out on the sensation of sitting on a cold, hard bench for hours.

An essential element of immersion is interactivity. Watching TV is typically a passive experience. Full-blown virtual reality, by contrast, will be interactive. You will be able to move about in a virtual world, pick up objects you see, and communicate with people you meet. (A real football experience crucially includes the possibility of shouting abuse at

the referee.) To enable interactivity, the system must have sensors that pick up on your movements and utterances and adjust the presentation to incorporate the consequences of your actions.

Virtual worlds can be modeled on physical realities. If you are participating in a remote event through VR, as in the example of the imagined football spectator, you are said to be *telepresent* at that event. Virtual environments can also be wholly artificial, like cartoons, and have no particular counterpart in physical reality. Another possibility, known as *augmented reality*, is to have your perception of your immediate surroundings partially overlaid with simulated elements. For example, by wearing special glasses, nametags could be made to appear over the heads of guests at a dinner party, or you could opt to have annoying billboard advertisements blotted out from your view.

Many users of today's VR systems experience "simulator sickness," with symptoms ranging from unpleasantness and disorientation to headaches, nausea, and vomiting. Simulator sickness arises because different sensory systems provide conflicting cues. For example, the visual system may provide strong cues of self-motion while the vestibular system in your inner ear tells your brain that your head is stationary. Heavy head-mounted display helmets and lag times between tracking device and graphics update can also cause discomfort. Creating good VR that overcomes these problems is technically challenging.

Primitive virtual realities have been around for some time. Early applications included training modules for pilots and military personnel. Increasingly, VR is used in computer gaming. Partly because VR is computationally very intensive, simulations are still quite crude. As computational power increases, and as sensors, effectors and displays improve, VR could begin to approximate physical reality in terms of fidelity and interactivity.

In the long run, VR could unlock limitless possibilities for human creativity. We could construct artificial experiential worlds, in which the laws of physics can be suspended, that would appear as real as physical reality to participants. People could visit these worlds for work, entertainment, or to socialize with friends who may be living on the opposite site of the globe. Uploads [see "What is uploading?"], who could interact with simulated environments directly without the need of a mechanical interface, might spend most of their time in virtual realities.

2.5 What is cryonics? Isn't the probability of success too small?

Cryonics is an experimental medical procedure that seeks to save lives by placing in low-temperature storage persons who cannot be treated with current medical procedures and who have been declared legally dead, in the hope that technological progress will eventually make it possible to revive them.

For cryonics to work today, it is not necessary that we can currently reanimate cryo-preserved patients (which we cannot). All that is needed is that we can preserve patients

in a state sufficiently intact that *some possible technology*, developed in the future, will one day be able to repair the freezing damage and reverse the original cause of deanimation. Only half of the complete cryonics procedure can be scrutinized today; the other half cannot be performed until the (perhaps distant) future.

What we know now is that it is possible to stabilize a patient's condition by cooling him or her in liquid nitrogen (- 196 C°). A considerable amount of cell damage is caused by the freezing process. This injury can be minimized by following suspension protocols that involve suffusing the deanimated body with cryoprotectants. The formation of damaging ice crystals can even be suppressed altogether in a process known as vitrification, in which the patient's body is turned into a kind of glass. This might sound like an improbable treatment, but the purpose of cryonics is to preserve the *structure* of life rather than the *processes* of life, because the life processes can in principle be re-started as long as the information encoded in the structural properties of the body, in particular the brain, are sufficiently preserved. Once frozen, the patient can be stored for millennia with virtually no further tissue degradation.

Many experts in molecular nanotechnology believe that in its mature stage nanotechnology will enable the revival of cryonics patients. Hence, it is possible that the suspended patients could be revived in as little as a few decades from now. The uncertainty about the ultimate technical feasibility of reanimation may very well be dwarfed by the uncertainty in other factors, such as the possibility that you deanimate in the wrong kind of way (by being lost at sea, for example, or by having the brain's information content erased by Alzheimer's disease), that your cryonics company goes bust, that civilization collapses, or that people in the future won't be interested in reviving you. So, a cryonics contract is far short of a survival guarantee. As a cryonicist saying goes, being cryonically suspended is the second worst thing that can happen to you.

When we consider the procedures that are routine today and how they might have been viewed in (say) the 1700s, we can begin to see how difficult it is to make a well-founded argument that future medical technology will never be able to reverse the injuries that occur during cryonic suspension. By contrast, your chances of a this-worldly comeback if you opt for one of the popular alternative treatments – such as cremation or burial – are zero. Seen in this light, signing up for cryonics, which is usually done by making a cryonics firm one of the beneficiaries of your life insurance, can look like a reasonable insurance policy. If it doesn't work, you would be dead anyway. If it works, it may save your life. Your saved life would then likely be extremely long and healthy, given how advanced the state of medicine must be to revive you.

By no means are all transhumanists signed up for cryonics, but a significant fraction finds that, for them, a cost-benefit analysis justifies the expense. Becoming a cryonicist, however, requires courage: the courage to confront the possibility of your own death, and the courage to resist the peer-pressure from the large portion of the population which currently espouses deathist values and advocates complacency in the face of a continual, massive loss of human life.

References:

Merkle, R. "The Molecular Repair of the Brain." *Cryonics magazine*, Vol. 15, No's 1 & 2. (1994).
<http://www.merkle.com/cryo/techFeas.html>

2.6 What is uploading?

Uploading (sometimes called "downloading", "mind uploading" or "brain reconstruction") is the process of transferring an intellect from a biological brain to a computer.

One way of doing this might be by first scanning the synaptic structure of a particular brain and then implementing the same computations in an electronic medium. A brain scan of sufficient resolution could be produced by disassembling the brain atom for atom by means of nanotechnology. Other approaches, such as analyzing pieces of the brain slice by slice in an electron microscope with automatic image processing have also been proposed. In addition to mapping the connection pattern among the 100 billion-or-so neurons, the scan would probably also have to register some of the functional properties of each of the synaptic interconnections, such as the efficacy of the connection and how stable it is over time (e.g. whether it is short-term or long-term potentiated). Non-local modulators such as neurotransmitter concentrations and hormone balances may also need to be represented, although such parameters likely contain much less data than the neuronal network itself.

In addition to a good three-dimensional map of a brain, uploading will require progress in neuroscience to develop functional models of each species of neuron (how they map input stimuli to outgoing action potentials, and how their properties change in response to activity in learning). It will also require a powerful computer to run the upload, and some way for the upload to interact with the external world or with a virtual reality. (Providing input/output or a virtual reality for the upload appears easy in comparison to the other challenges.)

An alternative hypothetical uploading method would proceed more gradually: one neuron could be replaced by an implant or by a simulation in a computer outside of the body. Then another neuron, and so on, until eventually the whole cortex has been replaced and the person's thinking is implemented on entirely artificial hardware. (To do this for the whole brain would almost certainly require nanotechnology.)

A distinction is sometimes made between destructive uploading, in which the original brain is destroyed in the process, and non-destructive uploading, in which the original brain is preserved intact alongside the uploaded copy. It is a matter of debate under what conditions personal identity would be preserved in destructive uploading. Many philosophers who have studied the problem think that at least under some conditions, an upload of your brain would be you. A widely accepted position is that you survive so long as certain information patterns are conserved, such as your memories, values, attitudes, and emotional dispositions, and so long as there is causal continuity so that earlier stages of yourself help determine later stages of yourself. Views differ on the

relative importance of these two criteria, but they can *both* be satisfied in the case of uploading. For the continuation of personhood, on this view, it matters little whether you are implemented on a silicon chip inside a computer or in that gray, cheesy lump inside your skull, assuming both implementations are conscious.

Tricky cases arise, however, if we imagine that several similar copies are made of your uploaded mind. Which one of them is you? Are they all you, or are none of them you? Who owns your property? Who is married to your spouse? Philosophical, legal, and ethical challenges abound. Maybe these will become hotly debated political issues later in this century.

A common misunderstanding about uploads is that they would necessarily be “disembodied” and that this would mean that their experiences would be impoverished. Uploading according to this view would be the ultimate escapism, one that only neurotic body-loathers could possibly feel tempted by. But an upload’s experience could in principle be identical to that of a biological human. An upload could have a virtual (simulated) body giving the same sensations and the same possibilities for interaction as a non-simulated body. With advanced virtual reality, uploads could enjoy food and drink, and upload sex could be as gloriously messy as one could wish. And uploads wouldn’t have to be confined to virtual reality: they could interact with people on the outside and even rent robot bodies in order to work in or explore physical reality.

Personal inclinations regarding uploading differ. Many transhumanists have a pragmatic attitude: whether they would like to upload or not depends on the precise conditions in which they would live as uploads and what the alternatives are. (Some transhumanists may also doubt whether uploading will be possible.) Advantages of being an upload would include:

- Uploads would not be subject to biological senescence.
- Back-up copies of uploads could be created regularly so that you could be re-booted if something bad happened. (Thus your lifespan would potentially be as long as the universe’s.)
- You could potentially live much more economically as an upload since you wouldn’t need physical food, housing, transportation, etc.
- If you were running on a fast computer, you would think faster than in a biological implementation. For instance, if you were running on a computer a thousand times more powerful than a human brain, then you would think a thousand times faster (and the external world would appear to you as if it were slowed down by a factor of a thousand). You would thus get to experience more subjective time, and live more, during any given day.
- You could travel at the speed of light as an information pattern, which could be convenient in a future age of large-scale space settlements.

- Radical cognitive enhancements would likely be easier to implement in an upload than in an organic brain.

A couple of other points about uploading:

- Uploading should work for cryonics patients provided their brains are preserved in a sufficiently intact state.
- Uploads could reproduce extremely quickly (simply by making copies of themselves). This implies that resources could very quickly become scarce unless reproduction is regulated.

2.7 What is the singularity?

Some thinkers conjecture that there will be a point in the future when the rate of technological development becomes so rapid that the progress-curve becomes nearly vertical. Within a very brief time (months, days, or even just hours), the world might be transformed almost beyond recognition. This hypothetical point is referred to as the singularity. The most likely cause of a singularity would be the creation of some form of rapidly self-enhancing greater-than-human intelligence.

The concept of the singularity is often associated with Vernor Vinge, who regards it as one of the more probable scenarios for the future. (Earlier intimations of the same idea can be found e.g. in John von Neumann, as paraphrased by Ulam 1958, and in I. J. Good 1965.) Provided that we manage to avoid destroying civilization, Vinge thinks that a singularity is likely to happen as a consequence of advances in artificial intelligence, large systems of networked computers, computer-human integration, or some other form of intelligence amplification. Enhancing intelligence will, in this scenario, at some point lead to a positive feedback loop: smarter systems can design systems that are even more intelligent, and can do so more swiftly than the original human designers. This positive feedback effect would be powerful enough to drive an intelligence explosion that could quickly lead to the emergence of a superintelligent system of surpassing abilities.

The singularity-hypothesis is sometimes paired with the claim that it is impossible for us to predict what comes after the singularity. A post-singularity society might be so alien that we can know nothing about it. One exception might be the basic laws of physics, but even there it is sometimes suggested that there may be undiscovered laws (for instance, we don't yet have an accepted theory of quantum gravity) or poorly understood consequences of known laws that could be exploited to enable things we would normally think of as physically impossible, such as creating traversable wormholes, spawning new "basement" universes, or traveling backward in time. However, unpredictability is logically distinct from abruptness of development and would need to be argued for separately.

Transhumanists differ widely in the probability they assign to Vinge's scenario. Almost all of those who do think that there will be a singularity believe it will happen in this century, and many think it is likely to happen within several decades.

References:

Good, I. J. "Speculations Concerning the First Ultraintelligent Machine," in *Advances in Computers*, Vol. 6, Franz L. Alt and Morris Rubinoff, eds (Academic Press, 1965), pp. 31-88.

Vinge, V. "The Coming Technological Singularity," *Whole Earth Review*, Winter Issue (1993).
<http://www.ugcs.caltech.edu/~phoenix/vinge/vinge-sing.html>

Ulam, S. "Tribute to John von Neumann," *Bulletin of the American Mathematical Society*, Vol. 64, Nr. 3, Part II, pp. 1-49 (1958).

3 SOCIETY AND POLITICS

3.1 Will new technologies only benefit the rich and powerful?

One could make the case that the average citizen of a developed country today has a higher standard of living than any king five hundred years ago. The king might have had a court orchestra, but you can afford a CD player that lets you to listen to the best musicians any time you want. When the king got pneumonia he might well die, but you can take antibiotics. The king might have a carriage with six white horses, but you can have a car that is faster and more comfortable. And you likely have television, Internet access, and a shower with warm water; you can talk with relatives who live in a different country over the phone; and you know more about the Earth, nature, and the cosmos than any medieval monarch.

The typical pattern with new technologies is that they become cheaper as time goes by. In the medical field, for example, experimental procedures are usually available only to research subjects and the very rich. As these procedures become routine, costs fall and more people can afford them. Even in the poorest countries, millions of people have benefited from vaccines and penicillin. In the field of consumer electronics, the price of computers and other devices that were cutting-edge only a couple of years ago drops precipitously as new models are introduced.

It is clear that everybody can benefit greatly from improved technology. Initially, however, the greatest advantages will go to those who have the resources, the skills, and the willingness to learn to use new tools. One can speculate that some technologies may cause social inequalities to widen. For example, if some form of intelligence amplification becomes available, it may at first be so expensive that only the wealthiest can afford it. The same could happen when we learn how to genetically enhance our children. Those who are already well off would become smarter and make even more money. This phenomenon is not new. Rich parents send their kids to better schools and

provide them with resources such as personal connections and information technology that may not be available to the less privileged. Such advantages lead to greater earnings later in life and serve to increase social inequalities.

Trying to ban technological innovation on these grounds, however, would be misguided. If a society judges existing inequalities to be unacceptable, a wiser remedy would be progressive taxation and the provision of community-funded services such as education, IT access in public libraries, genetic enhancements covered by social security, and so forth. Economic and technological progress is not a zero sum game; it's a positive sum game. Technological progress does not solve the hard old political problem of what degree of income redistribution is desirable, but it can greatly increase the size of the pie that is to be divided.

3.2 Do transhumanists advocate eugenics?

Eugenics in the narrow sense refers to the pre-WWII movement in Europe and the United States to involuntarily sterilize the “genetically unfit” and encourage breeding of the genetically advantaged. These ideas are entirely contrary to the tolerant humanistic and scientific tenets of transhumanism. In addition to condemning the coercion involved in such policies, transhumanists strongly reject the racist and classist assumptions on which they were based, along with the notion that eugenic improvements could be accomplished in a practically meaningful timeframe through selective human breeding.

Transhumanists uphold the principles of bodily autonomy and procreative liberty. Parents must be allowed to choose for themselves whether to reproduce, how to reproduce, and what technological methods they use in their reproduction. The use of genetic medicine or embryonic screening to increase the probability of a healthy, happy, and multiply talented child is a responsible and justifiable application of parental reproductive freedom.

Beyond this, one can argue that parents have a moral responsibility to make use of these methods, assuming they are safe and effective. Just as it would be wrong for parents to fail in their duty to procure the best available medical care for their sick child, it would be wrong not to take reasonable precautions to ensure that a child-to-be will be as healthy as possible. This, however, is a moral judgment that is best left to individual conscience rather than imposed by law. Only in extreme and unusual cases might state infringement of procreative liberty be justified. If, for example, a would-be parent wished to undertake a genetic modification that would be clearly harmful to the child or would drastically curtail its options in life, then this prospective parent should be prevented by law from doing so. This case is analogous to the state taking custody of a child in situations of gross parental neglect or child abuse.

This defense of procreative liberty is compatible with the view that states and charities can subsidize public health, prenatal care, genetic counseling, contraception, abortion, and genetic therapies so that parents can make free and informed reproductive decisions

that result in fewer disabilities in the next generation. Some disability activists would call these policies eugenic, but society may have a legitimate interest in whether children are born healthy or disabled, leading it to subsidize the birth of healthy children, without actually outlawing or imposing particular genetic modifications.

When discussing the morality of genetic enhancements, it is useful to be aware of the distinction between enhancements that are intrinsically beneficial to the child or society on the one hand, and, on the other, enhancements that provide a merely positional advantage to the child. For example, health, cognitive abilities, and emotional well-being are valued by most people for their own sake. It is simply nice to be healthy, happy and to be able to think well, quite independently of any other advantages that come from possessing these attributes. By contrast, traits such as attractiveness, athletic prowess, height, and assertiveness seem to confer benefits that are mostly positional, i.e. they benefit a person by making her more competitive (e.g. in sports or as a potential mate), at the expense of those with whom she will compete, who suffer a corresponding disadvantage from her enhancement. Enhancements that have only positional advantages ought to be de-emphasized, while enhancements that create net benefits ought to be encouraged.

It is sometimes claimed that the use of germinal choice technologies would lead to an undesirable uniformity of the population. Some degree of uniformity is desirable and expected if we are able to make everyone congenitally healthy, strong, intelligent, and attractive. Few would argue that we should preserve cystic fibrosis because of its contribution to diversity. But other kinds of diversity are sure to flourish in a society with germinal choice, especially once adults are able to adapt their own bodies according to their own aesthetic tastes. Presumably most Asian parents will still choose to have children with Asian features, and if some parents choose genes that encourage athleticism, others may choose genes that correlate with musical ability.

It is unlikely that germ-line genetic enhancements will ever have a large impact on the world. It will take a minimum of forty or fifty years for the requisite technologies to be developed, tested, and widely applied and for a significant number of enhanced individuals to be born and reach adulthood. Before this happens, more powerful and direct methods for individuals to enhance themselves will probably be available, based on nanomedicine, artificial intelligence, uploading, or somatic gene therapy. (Traditional eugenics, based on selecting who is allowed to reproduce, would have even less prospect of avoiding preemptive obsolescence, as it would take many generations to deliver its purported improvements.)

3.3 Aren't these future technologies very risky? Could they even cause our extinction?

Yes, and this implies an urgent need to analyze the risks before they materialize and to take steps to reduce them. Biotechnology, nanotechnology, and artificial intelligence pose

especially serious risks of accidents and abuse. [See also “If these technologies are so dangerous, should they be banned? What can be done to reduce the risks?”]

One can distinguish between, on the one hand, endurable or limited hazards, such as car crashes, nuclear reactor meltdowns, carcinogenic pollutants in the atmosphere, floods, volcano eruptions, and so forth, and, on the other hand, *existential risks* – events that would cause the extinction of intelligent life or permanently and drastically cripple its potential. While endurable or limited risks can be serious – and may indeed be fatal to the people immediately exposed – they are recoverable; they do not destroy the long-term prospects of humanity as a whole. Humanity has long experience with endurable risks and a variety of institutional and technological mechanisms have been employed to reduce their incidence. Existential risks are a different kind of beast. For most of human history, there were no significant existential risks, or at least none that our ancestors could do anything about. By definition, of course, no existential disaster has yet happened. As a species we may therefore be less well prepared to understand and manage this new kind of risk. Furthermore, the reduction of existential risk is a global public good (everybody by necessity benefits from such safety measures, whether or not they contribute to their development), creating a potential free-rider problem, i.e. a lack of sufficient selfish incentives for people to make sacrifices to reduce an existential risk. Transhumanists therefore recognize a moral duty to promote efforts to reduce existential risks.

The gravest existential risks facing us in the coming decades will be of our own making. These include:

Destructive uses of nanotechnology. The accidental release of a self-replicating nanobot into the environment, where it would proceed to destroy the entire biosphere, is known as the “gray goo scenario”. Since molecular nanotechnology will make use of positional assembly to create non-biological structures and to open new chemical reaction pathways, there is no reason to suppose that the ecological checks and balances that limit the proliferation of organic self-replicators would also contain nano-replicators. Yet, while gray goo is certainly a legitimate concern, relatively simple engineering safeguards have been described that would make the probability of such a mishap almost arbitrarily small (Foresight 2002). Much more serious is the threat posed by nanobots deliberately designed to be destructive. A terrorist group or even a lone psychopath, having obtained access to this technology, could do extensive damage or even annihilate life on earth unless effective defensive technologies had been developed beforehand (Center for Responsible Nanotechnology 2003). An unstable arms race between nanotechnic states could also result in our eventual demise (Gubrud 2000). Anti-proliferation efforts will be complicated by the fact that nanotechnology does not require difficult-to-obtain raw materials or large manufacturing plants, and by the dual-use functionality of many of the basic components of destructive nanomachinery. While a nanotechnic defense system (which would act as a global immune system capable of identifying and neutralizing rogue replicators) appears to be possible in principle, it could turn out to be more difficult to construct than a simple destructive replicator. This could create a window of global vulnerability between the potential creation of dangerous replicators and the development

of an effective immune system. It is critical that nano-assemblers do not fall into the wrong hands during this period.

Biological warfare. Progress in genetic engineering will lead not only to improvements in medicine but also to the capability to create more effective bioweapons. It is chilling to consider what would have happened if HIV had been as contagious as the virus that causes the common cold. Engineering such microbes might soon become possible for increasing numbers of people. If the RNA sequence of a virus is posted on the Internet, then anybody with some basic expertise and access to a lab will be able to synthesize the actual virus from this description. A demonstration of this possibility was offered by a small team of researchers from New York University at Stony Brook in 2002, who synthesized the polio virus (whose genetic sequence is on the Internet) from scratch and injected it into mice who subsequently became paralyzed and died.

Artificial intelligence. No threat to human existence is posed by today's AI systems or their near-term successors. But if and when superintelligence is created, it will be of paramount importance that it be endowed with human-friendly values. An imprudently or maliciously designed superintelligence, with goals amounting to indifference or hostility to human welfare, could cause our extinction. Another concern is that the first superintelligence, which may become very powerful because of its superior planning ability and because of the technologies it could swiftly develop, would be built to serve only a single person or a small group (such as its programmers or the corporation that commissioned it). While this scenario may not entail the extinction of literally all intelligent life, it nevertheless constitutes an existential risk because the future that would result would be one in which a great part of humanity's potential had been permanently destroyed and in which at most a tiny fraction of all humans would get to enjoy the benefits of posthumanity. [See also "Will posthumans or superintelligent machines pose a threat to humans who aren't augmented?"]

Nuclear war. Today's nuclear arsenals are probably not sufficient to cause the extinction of all humans, but future arms races could result in even larger build-ups. It is also conceivable that an all-out nuclear war would lead to the collapse of modern civilization, and it is not completely certain that the survivors would succeed in rebuilding a civilization capable of sustaining growth and technological development.

Something unknown. All the above risks were unknown a century ago and several of them have only become clearly understood in the past two decades. It is possible that there are future threats of which we haven't yet become aware.

For a more extensive discussion of these and many other existential risks, see Bostrom (2002).

Evaluating the total probability that some existential disaster will do us in before we get the opportunity to become posthuman can be done by various direct or indirect methods. Although any estimate inevitably includes a large subjective factor, it seems that to set the probability to less than 20% would be unduly optimistic, and the best estimate may be

considerably higher. But depending on the actions we take, this figure can be raised or lowered.

References:

Bostrom, N. "Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards," *Journal of Evolution and Technology*. Vol. 9 (2002). <http://www.nickbostrom.com/existential/risks.html>

Center for Responsible Nanotechnology. "Dangers of Nanotechnology" (2003). <http://www.crnano.org/dangers.htm>

Foresight Institute. "Foresight Guidelines on Molecular Nanotechnology, version 3.7" (2000). <http://www.foresight.org/guidelines/current.html>

Gubrud, M. "Nanotechnology and International Security," *Fifth Foresight Conference on Molecular Nanotechnology*. (1997) <http://www.foresight.org/Conferences/MNT05/Papers/Gubrud/index.html>

Wimmer, E. et al. "Chemical Synthesis of Poliovirus cDNA: Generation of Infectious Virus in the Absence of Natural Template," *Science*, Vol. 257, No. 5583, (2002), pp. 1016-1018.

3.4 If these technologies are so dangerous, should they be banned? What can be done to reduce the risks?

The position that we ought to relinquish research into robotics, genetic engineering, and nanotechnology has been advocated in an article by Bill Joy (2000). Joy argued that some of the future applications of these technologies are so dangerous that research in those fields should be stopped now. Partly because of Joy's previously technophilic credentials (he was a software designer and a cofounder of Sun Microsystems), his article, which appeared in *Wired* magazine, attracted a great deal of attention.

Many of the responses to Joy's article pointed out that there is no realistic prospect of a worldwide ban on these technologies; that they have enormous potential benefits that we would not want to forgo; that the poorest people may have a higher tolerance for risk in developments that could improve their condition; and that a ban may actually increase the dangers rather than reduce them, both by delaying the development of protective applications of these technologies, and by weakening the position of those who choose to comply with the ban relative to less scrupulous groups who defy it.

A more promising alternative than a blanket ban is *differential technological development*, in which we would seek to influence the *sequence* in which technologies developed. On this approach, we would strive to retard the development of harmful technologies and their applications, while accelerating the development of beneficial technologies, especially those that offer protection against the harmful ones. For technologies that have decisive military applications, unless they can be verifiably banned, we may seek to ensure that they are developed at a faster pace in countries we regard as responsible than in those that we see as potential enemies. (Whether a ban is verifiable and enforceable can change over time as a result of developments in the international system or in surveillance technology.)

In the case of nanotechnology, the desirable sequence of development is that nanotech immune systems and other defensive measures be deployed before offensive capabilities become available to many independent powers. Once a technology is shared by many, it becomes extremely hard to prevent further proliferation. In the case of biotechnology, we should seek to promote research into vaccines, anti-viral drugs, protective gear, sensors, and diagnostics, and to delay as long as possible the development and proliferation of biological warfare agents and the means of their weaponization. For artificial intelligence, a serious risk will emerge only when capabilities approach or surpass those of humans. At that point one should seek to promote the development of friendly AI and to prevent unfriendly or unreliable AI systems.

Superintelligence is an example of a technology that seems especially worth promoting because it can help reduce a broad range of threats. Superintelligent systems could advise us on policy and make the progress curve for nanotechnology steeper, thus shortening the period of vulnerability between the development of dangerous nanoreplicators and the deployment of effective defenses. If we have a choice, it seems preferable that superintelligence be developed before advanced nanotechnology, as superintelligence could help reduce the risks of nanotechnology but not vice versa. Other technologies that have wide risk-reducing uses include intelligence augmentation, information technology, and surveillance. These can make us smarter individually and collectively or make enforcement of necessary regulation more feasible. A strong *prima facie* case therefore exists for pursuing these technologies as vigorously as possible. Needless to say, we should also promote non-technological developments that are beneficial in almost all scenarios, such as peace and international cooperation.

In confronting the hydra of existential, limited, and endurable risks glaring at us from the future, it is unlikely that any one silver bullet will provide adequate protection. Instead, an arsenal of countermeasures will be needed so that we can address the various risks on multiple levels.

The first step to tackling a risk is to recognize its existence. More research is needed, and existential risks in particular should be singled out for attention because of their seriousness and because of the special nature of the challenges they pose. Surprisingly little work has been done in this area (but see e.g. Leslie (1996), Bostrom (2002), and Rees (2003) for some preliminary explorations). The strategic dimensions of our choices must be taken into account, given that some of the technologies in questions have important military ramifications. In addition to scholarly studies of the threats and their possible countermeasures, public awareness must be raised to enable a more informed debate of our long-term options.

Some of the lesser existential risks, such as an apocalyptic asteroid impact or the highly speculative scenario involving something like the upsetting of a metastable vacuum state in some future particle accelerator experiment, could be substantially reduced at relatively small expense. Programs to accomplish this – e.g. an early detection system for dangerous near-earth objects on potential collision course with Earth, or the

commissioning of advance peer review of planned high-energy physics experiments – are probably cost-effective. However, these lesser risks must not deflect attention from the more serious concern raised by more probable existential disasters [see “Aren’t these future technologies very risky? Could they even cause our extinction?”].

In light of how superabundant the human benefits of technology can ultimately be, it matters less that we obtain all of these benefits in their precisely most optimal form, and more that we obtain them at all. For many practical purposes, it makes sense to adopt the rule of thumb that we should act so as to maximize the probability of an *acceptable* outcome, one in which we attain some (reasonably broad) realization of our potential; or, to put it in negative terms, that we should act so as to minimize net existential risk.

References:

Bostrom, N. “Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards,” *Journal of Evolution and Technology*. Vol. 9 (2002). <http://www.nickbostrom.com/existential/risks.html>

Joy, B. “Why the Future Doesn’t Need Us”. *Wired*, 8:04 (2000).
http://www.wired.com/wired/archive/8.04/joy_pr.html

Leslie, J. *The End of the World: The Ethics and Science of Human Extinction*. (London: Routledge, 1996).

Rees, M. *Our Final Hour*. (New York: Basic Books, 2003).

3.5 Shouldn’t we concentrate on current problems such as improving the situation of the poor, rather than putting our efforts into planning for the “far” future?

We should do both. Focusing solely on current problems would leave us unprepared for the new challenges that we will encounter.

Many of the technologies and trends that transhumanists discuss are already reality. Biotechnology and information technology have transformed large sectors of our economies. The relevance of transhumanist ethics is manifest in such contemporary issues as stem cell research, genetically modified crops, human genetic therapy, embryo screening, end of life decisions, enhancement medicine, information markets, and research funding priorities. The importance of transhumanist ideas is likely to increase as the opportunities for human enhancement proliferate.

Transhuman technologies will tend to work well together and create synergies with other parts of human society. For example, one important factor in healthy life expectancy is access to good medical care. Improvements in medical care will extend healthy, active lifespan – “healthspan” – and research into healthspan extension is likely to benefit ordinary care. Work on amplifying intelligence has obvious applications in education, decision-making, and communication. Better communications would facilitate trade and understanding between people. As more and more people get access to the Internet and are able to receive satellite radio and television broadcasts, dictators and totalitarian

regimes may find it harder to silence voices of dissent and to control the information flow in their populations. And with the Internet and email, people discover they can easily form friendships and business partnerships in foreign countries. A world order characterized by peace, international cooperation, and respect for human rights would much improve the odds that the potentially dangerous applications of some future technologies can be controlled and would also free up resources currently spent on military armaments, some of which could then hopefully be diverted to improving the condition of the poor. Nanotechnological manufacturing promises to be both economically profitable and environmentally sound. Transhumanists do not have a patent solution to achieve these outcomes, any more than anybody else has, but technology has a huge role to play.

An argument can be made that the most efficient way of contributing to making the world better is by participating in the transhumanist project. This is so because the stakes are enormous – humanity's entire future may depend on how we manage the coming technological transitions – and because relatively few resources are at the present time being devoted to transhumanist efforts. Even one extra person can still make a significant difference here.

3.6 Will extended life worsen overpopulation problems?

Population increase is an issue we would ultimately have to come to grips with even if healthy life-extension were not to happen. Leaving people to die is an unacceptable solution.

A large population should not be viewed simply as a problem. Another way of looking at the same fact is that it means that many persons now enjoy lives that would not have been lived if the population had been smaller. One could ask those who complain about overpopulation exactly which people's lives they would have preferred should not have been led. Would it really have been better if billions of the world's people had never existed and if there had been no other people in their place? Of course, this is not to deny that too-rapid population growth can cause crowding, poverty, and the depletion of natural resources. In this sense there can be real problems that need to be tackled.

How many people the Earth can sustain at a comfortable standard of living is a function of technological development (as well as of how resources are distributed). New technologies, from simple improvements in irrigation and management, to better mining techniques and more efficient power generation machinery, to genetically engineered crops, can continue to improve world resource and food output, while at the same time reducing environmental impact and animal suffering.

Environmentalists are right to insist that the status quo is unsustainable. As a matter of physical necessity, things cannot stay as they are today indefinitely, or even for very long. If we continue to use up resources at the current pace, without finding more resources or learning how to use novel kinds of resources, then we will run into serious

shortages sometime around the middle of this century. The deep greens have an answer to this: they suggest we turn back the clock and return to an idyllic pre-industrial age to live in sustainable harmony with nature. The problem with this view is that the pre-industrial age was anything but idyllic. It was a life of poverty, misery, disease, heavy manual toil from dawn to dusk, superstitious fears, and cultural parochialism. Nor was it environmentally sound – as witness the deforestation of England and the Mediterranean region, desertification of large parts of the middle east, soil depletion by the Anasazi in the Glen Canyon area, destruction of farm land in ancient Mesopotamia through the accumulation of mineral salts from irrigation, deforestation and consequent soil erosion by the ancient Mexican Mayas, overhunting of big game almost everywhere, and the extinction of the dodo and other big featherless birds in the South Pacific. Furthermore, it is hard to see how more than a few hundred million people could be maintained at a reasonable standard of living with pre-industrial production methods, so some ninety percent of the world population would somehow have to vanish in order to facilitate this nostalgic return.

Transhumanists propose a much more realistic alternative: not to retreat to an imagined past, but to press ahead as intelligently as we can. The environmental problems that technology creates are problems of intermediary, inefficient technology, of placing insufficient political priority on environmental protection as well as of a lack of ecological knowledge. Technologically less advanced industries in the former Soviet-bloc pollute much more than do their advanced Western counterparts. High-tech industry is typically relatively benign. Once we develop molecular nanotechnology, we will not only have clean and efficient manufacturing of almost any commodity, but we will also be able to clean up much of the mess created by today's crude fabrication methods. This would set a standard for a clean environment that today's traditional environmentalists could scarcely dream of.

Nanotechnology will also make it cheaper to colonize space. From a cosmic point of view, Earth is an insignificant speck. It has sometimes been suggested that we ought to leave space untouched in its pristine glory. This view is hard to take seriously. Every hour, through entirely natural processes, vast amounts of resources – millions of times more than the sum total of what the human species has consumed throughout its career – are transformed into radioactive substances or wasted as radiation escaping into intergalactic space. Can we not think of some more creative way of using all this matter and energy?

Even with full-blown space colonization, however, population growth can continue to be a problem, and this is so even if we assume that an unlimited number of people could be transported from Earth into space. If the speed of light provides an upper bound on the expansion speed then the amount of resources under human control will grow only polynomially ($\sim t^3$). Population, on the other hand, can easily grow exponentially ($\sim e^t$). If that happens, then, since a factor that grows exponentially will eventually overtake any factor that grows polynomially, average income will ultimately drop to subsistence levels, forcing population growth to slow. How soon this would happen depends primarily on reproduction rates. A change in average life span would not have a big

effect. Even vastly improved technology can only postpone this inevitability for a relatively brief time. The only long-term method of assuring continued growth of average income is some form of population control, whether spontaneous or imposed, limiting the number of new persons created per year. This does not mean that population could not grow, only that the growth would have to be polynomial rather than exponential.

Some additional points to consider:

- In technologically advanced countries, couples tend to have fewer children, often below the replacement rate. As an empirical generalization, giving people increased rational control over their lives, especially through women's education and participation in the labor market, causes couples to have fewer children.
- If one took seriously the idea of controlling population by limiting life span, why not be more active about it? Why not encourage suicide? Why not execute anyone reaching the age of 75?
- If slowing aging were unacceptable because it might lead to there being more people, what about efforts to cure cancer, reduce traffic deaths, or improve worker safety? Why use double standards?
- When transhumanists say they want to extend lifespans, what they mean is that they want to extend healthspans. This means that the extra person-years would be productive and would add economic value to society. We can all agree that there would be little point in living an extra ten years in a state of dementia.
- The world population growth rate has been declining for several decades. It peaked in 1970 at 2.1%. In 2003, it was 1.2%; and it is expected to fall below 1.0% around 2015. (United Nations 2002). The doomsday predictions of the so-called "Club of Rome" from the early 1970s have consistently turned out to be wrong.
- The more people there are, the more brains there will be working to invent new ideas and solutions.
- If people can look forward to a longer healthy, active life, they will have a personal stake in the future and will hopefully be more concerned about the long-term consequences of their actions.

References:

United Nations. *The World Population Prospects: The 2002 Revision* (United Nations: New York, 2002).
<http://www.gov.za/reports/2003/unpdhighlights.pdf>

3.7 Is there any ethical standard by which transhumanists judge “improvement of the human condition”?

Transhumanism is compatible with a variety of ethical systems, and transhumanists themselves hold many different views. Nonetheless, the following seems to constitute a common core of agreement:

According to transhumanists, the human condition has been improved if the conditions of individual humans have been improved. In practice, competent adults are usually the best judges of what is good for themselves. Therefore, transhumanists advocate individual freedom, especially the right for those who so wish to use technology to extend their mental and physical capacities and to improve their control over their own lives.

From this perspective, an improvement to the human condition is a change that gives increased opportunity for individuals to shape themselves and their lives according to their informed wishes. Notice the word “informed”. It is important that people be aware of what they choose between. Education, discussion, public debate, critical thinking, artistic exploration, and, potentially, cognitive enhancers are means that can help people make more informed choices.

Transhumanists hold that people are not disposable. Saving lives (of those who want to live) is ethically important. It would be wrong to unnecessarily let existing people die in order to replace them with some new “better” people. Healthspan-extension and cryonics are therefore high on the transhumanist list of priorities. The transhumanist goal is not to replace existing humans with a new breed of super-beings, but rather to give human beings (those existing today and those who will be born in the future) the option of developing into posthuman persons.

The non-disposability of persons partially accounts for a certain sense of urgency that is common among transhumanists. On average, 150,000 men, women, and children die every day, often in miserable conditions. In order to give as many people as possible the chance of a posthuman existence – or even just a decent human existence – it is paramount that technological development, in at least some fields, is pursued with maximal speed. When it comes to life-extension and its various enabling technologies, a delay of a single week equals one million avoidable premature deaths – a weighty fact which those who argue for bans or moratoria would do well to consider carefully. (The further fact that universal access will likely lag initial availability only adds to the reason for trying to hurry things along.)

Transhumanists reject speciesism, the (human racist) view that moral status is strongly tied to membership in a particular biological species, in our case *homo sapiens*. What exactly does determine moral status is a matter of debate. Factors such as being a person, being sentient, having the capacity for autonomous moral choice, or perhaps even being a member of the same community as the evaluator, are among the criteria that may combine to determine the degree of somebody’s moral status (Warren 1997). But transhumanists argue that species-identity should be de-emphasized in this context. Transhumanists insist that all beings that can experience pain have some moral status,

and that posthuman persons could have at least the same level of moral status as humans have in their current form.

References:

Warren, M.-A. *Moral Status: Obligations to Persons and Other Living Things* (Oxford: Oxford University Press, 1997).

3.8 What kind of society would posthumans live in?

Not enough information is available at the current time to provide a full answer to this question. In part, though, the answer is, “You decide.” The outcome may be influenced by the choices we make now and over the coming decades. In this respect, the situation is the same as in earlier epochs that had no transhuman possibilities: by becoming involved in political struggles against today’s social ills and injustices, we can help make tomorrow’s society better.

Transhumanism does, however, inform us about new constraints, possibilities, and issues, and it highlights numerous important leverage points for intervention, where a small application of resources can make a big long-term difference. For example, one issue that moves into prominence is the challenge of creating a society in which beings with vastly different orders of capabilities (such as posthuman persons and as-yet non-augmented humans) can live happily and peacefully together. Another concern that becomes paramount is the need to build a world order in which dangerous arms races can be prevented and in which the proliferation of weapons of mass destruction can be suppressed or at least delayed until effective defenses have been developed [see “Aren’t these future technologies very risky? Could they even cause our extinction?”].

The ideal social organization may be one that includes the possibility for those who so wish to form independent societies voluntarily secluded from the rest of the world, in order to pursue traditional ways of life or to experiment with new forms of communal living. Achieving an acceptable balance between the rights of such communities for autonomy, on the one hand, and the security concerns of outside entities and the just demands for protection of vulnerable and oppressed individuals inside these communities on the other hand, is a delicate task and a familiar challenge in political philosophy.

What types of society posthumans will live in depends on what types of posthumans eventually develop. One can project various possible developmental paths [see “What is a posthuman?”] which may result in very different kinds of posthuman, transhuman, and unaugmented human beings, living in very different sorts of societies. In attempting to imagine such a world, we must bear in mind that we are likely to base our expectations on the experiences, desires, and psychological characteristics of humans. Many of these expectations may not hold true of posthuman persons. When human nature changes, new ways of organizing a society may become feasible. We may hope to form a clearer understanding of what those new possibilities are as we observe the seeds of transhumanity develop.

3.9 Will posthumans or superintelligent machines pose a threat to humans who aren't augmented?

Human society is always at risk from some group deciding to view another group of humans as fit for slavery or slaughter. To counteract such tendencies, modern societies have created laws and institutions, and endowed them with powers of enforcement, that act to prevent groups of citizens from assaulting one another. The efficacy of these institutions does not depend on all citizens having equal capacities. Modern, peaceful societies have large numbers of people with diminished physical or mental capacities along with many other people who may be exceptionally physically strong or healthy or intellectually talented in various ways. Adding people with technologically enhanced capacities to this already broad distribution of ability would not necessarily rip society apart or trigger genocide or enslavement.

A common worry is that inheritable genetic modifications or other human enhancement technologies would lead to two distinct and separate species and that hostilities would inevitably develop between them. The assumptions behind this prediction should be questioned. It is a common theme in fiction because of the opportunities for dramatic conflict, but that is not the same as social, political, and economic plausibility in the real world. It seems more likely that there would be a continuum of differently modified or enhanced individuals, which would overlap with the continuum of as-yet unenhanced humans. The scenario in which “the enhanced” form a pact and then attack “the naturals” makes for exciting science fiction but is not necessarily the most plausible outcome. Even today, the segment containing the tallest 90 percent of the population could, in principle, get together and kill or enslave the shorter decile. That this does not happen suggests that a well-organized society can hold together even if it contains many possible coalitions of people sharing some attribute such that, if they unified under one banner, would make them capable of exterminating the rest.

To note that the extreme case of a war between human and posthuman persons is not the most likely scenario is not to say that there are no legitimate social concerns about the steps that may take us closer to posthumanity. Inequity, discrimination, and stigmatization – against or on behalf of modified people – could become serious issues. Transhumanists would argue that these (potential) social problems call for social remedies. (One case study of how contemporary technology can change important aspects of someone's identity is sex reassignment. The experiences of transsexuals show that some cultures still have work to do in becoming more accepting of diversity.) This is a task that we can begin to tackle now by fostering a climate of tolerance and acceptance towards those who are different from ourselves. We can also act to strengthen those institutions that prevent violence and protect human rights, for instance by building stable democratic traditions and constitutions and by expanding the rule of law to the international plane.

What about the hypothetical case in which someone intends to create, or turn themselves into, a being of so radically enhanced capacities that a single one or a small group of such

individuals would be capable of taking over the planet? This is clearly not a situation that is likely to arise in the imminent future, but one can imagine that, perhaps in a few decades, the prospective creation of superintelligent machines could raise this kind of concern. The would-be creator of a new life form with such surpassing capabilities would have an obligation to ensure that the proposed being is free from psychopathic tendencies and, more generally, that it has humane inclinations. For example, a superintelligence should be built with a clear goal structure that has friendliness to humans as its top goal. Before running such a program, the builders of a superintelligence should be required to make a strong case that launching it would be safer than alternative courses of action.

References:

Yudkowsky, E. *Creating Friendly AI: The Analysis and Design of Benevolent Goal Architectures*. (2003, Version 1.0). <http://www.singinst.org/CFAI/index.html>

4 TRANSHUMANISM AND NATURE

4.1 Why do transhumanists want to live longer?

This is a personal matter, a matter of the heart. Have you ever been so happy that you felt like melting into tears? Has there been a moment in your life of such depth and sublimity that the rest of existence seemed like dull, gray slumber from which you had only just woken up?

It is so easy to forget how good things can be when they are at their best. But on those occasions when we do remember – whether it comes from the total fulfillment of being immersed in creative work or from the tender ecstasy of reciprocated love – then we realize just how valuable every single minute of existence can be, when it is this good. And you might have thought to yourself, “It ought to be like this always. Why can’t this last forever?”

Well, maybe – just maybe – it could.

When transhumanists seek to extend human life, they are not trying to add a couple of extra years at a care home spent drooling at one’s shoes. The goal is more healthy, happy, productive years. Ideally, everybody should have the right to choose when and how to die – or not to die. Transhumanists want to live longer because they want to do, learn, and experience more; have more fun and spend more time with loved ones; continue to grow and mature beyond the paltry eight decades allotted to us by our evolutionary past; and in order to get to see for themselves what wonders the future might hold. As the sales pitch for one cryonics organization goes:

“The conduct of life and the wisdom of the heart are based upon time; in the last quartets of Beethoven, the last words and works of ‘old men’ like Sophocles and Russell and

Shaw, we see glimpses of a maturity and substance, an experience and understanding, a grace and a humanity, that isn't present in children or in teenagers. They attained it because they lived long; because they had time to experience and develop and reflect; time that we might all have. Imagine such individuals – a Benjamin Franklin, a Lincoln, a Newton, a Shakespeare, a Goethe, an Einstein [and a Gandhi] – enriching our world not for a few decades but for centuries. Imagine a world made of such individuals. It would truly be what Arthur C. Clarke called 'Childhood's End' – the beginning of the adulthood of humanity." (Cryonics Institute)

References:

Cryonics Institute. <http://www.cryonics.org/>

4.2 Isn't this tampering with nature?

Absolutely, and it is nothing to be ashamed of. It is often right to tamper with nature. One could say that manipulating nature is an important part of what civilization and human intelligence is all about; we have been doing it since the invention of the wheel.

Alternatively, one could say that since we are part of nature, everything we do and create is in a sense natural too. In any case, there is no moral reason why we shouldn't intervene in nature and improve it if we can, whether by eradicating diseases, improving agricultural yields to feed a growing world population, putting communication satellites into orbit to provide homes with news and entertainment, or inserting contact lenses in our eyes so we can see better. Changing nature for the better is a noble and glorious thing for humans to do. (On the other hand, to "pave paradise to put up a parking lot" would not be glorious; the qualification "for the better" is essential.) [See also "Are transhumanist technologies environmentally sound?"]

In many particular cases, of course, there are sound practical reasons for relying on "natural" processes. The point is that we cannot decide whether something is good or bad simply by asking whether it is natural or not. Some natural things are bad, such as starvation, polio, and being eaten alive by intestinal parasites. Some artificial things are bad, such as DDT-poisoning, car accidents, and nuclear war.

To pick a topical example, consider the debate about human cloning. Some argue that cloning humans is not unnatural because human clones are essentially just identical twins. They were right in this, of course, although one could also correctly remark that it is not natural for identical twins to be of different ages. But the more fundamental point is that it doesn't matter whether human clones are natural or not. When thinking about whether to permit human reproductive cloning, we have to compare the various possible desirable consequences with the various possible undesirable consequences. We then have to try to estimate the likelihood of each of these consequences. This kind of deliberation is much harder than simply dismissing cloning as unnatural, but it is also more likely to result in good decisions.

These remarks hopefully should seem trivial. Yet it is astonishing how often polemicists can still get a way with arguments that are basically (thinly disguised) ways of saying, "It

is good because it's the way it has always been!" or "It is good because that's the way Nature made it!"

4.3 Will transhuman technologies make us inhuman?

The important thing is not to be human but to be humane. Though we might wish to believe that Hitler was an inhuman monster, he was, in fact, a human monster; and Gandhi is noted not for being remarkably human but for being remarkably humane.

The attributes of our species are not exempt from ethical examination in virtue of being "natural" or "human". Some human attributes, such as empathy and a sense of fairness, are positive; others, such as tendencies toward tribalism or groupishness, have left deep scars on human history. If there is value in being human, it does not come from being "normal" or "natural", but from having within us the raw material for being humane: compassion, a sense of humor, curiosity, the wish to be a better person. Trying to preserve "humanness," rather than cultivating humaneness, would idolize the bad along with the good. One might say that if "human" is what we are, then "humane" is what we, as humans, wish we were. Human nature is not a bad place to start that journey, but we can't fulfill that potential if we reject any progress past the starting point.

4.4 Isn't death part of the natural order of things?

Transhumanists insist that whether something is natural or not is irrelevant to whether it is good or desirable [see also "Isn't this tampering with nature?", "Will extended life worsen overpopulation problems?", and "Why do transhumanists want to live longer?"].

Average human life span hovered between 20 and 30 years for most of our species' history. Most people today are thus living highly unnaturally long lives. Because of the high incidence of infectious disease, accidents, starvation, and violent death among our ancestors, very few of them lived much beyond 60 or 70. There was therefore little selection pressure to evolve the cellular repair mechanisms (and pay their metabolic costs) that would be required to keep us going beyond our meager three scores and ten. As a result of these circumstances in the distant past, we now suffer the inevitable decline of old age: damage accumulates at a faster pace than it can be repaired; tissues and organs begin to malfunction; and then we keel over and die.

The quest for immortality is one of the most ancient and deep-rooted of human aspirations. It has been an important theme in human literature from the very earliest preserved written story, *The Epic of Gilgamesh*, and in innumerable narratives and myths ever since. It underlies the teachings of world religions about spiritual immortality and the hope of an afterlife. If death is part of the natural order, so too is the human desire to overcome death.

Before transhumanism, the only hope of evading death was through reincarnation or otherworldly resurrection. Those who viewed such religious doctrines as figments of our own imagination had no alternative but to accept death as an inevitable fact of our existence. Secular worldviews, including traditional humanism, would typically include some sort of explanation of why death was not such a bad thing after all. Some existentialists even went so far as to maintain that death was necessary to give life meaning!

That people should make excuses for death is understandable. Until recently there was absolutely nothing anybody could do about it, and it made some degree of sense then to create comforting philosophies according to which dying of old age is a fine thing (“deathism”). If such beliefs were once relatively harmless, and perhaps even provided some therapeutic benefit, they have now outlived their purpose. Today, we can foresee the possibility of eventually abolishing aging and we have the option of taking active measures to stay alive until then, through life extension techniques and, as a last resort, cryonics. This makes the illusions of deathist philosophies dangerous, indeed fatal, since they teach helplessness and encourage passivity.

Espousing a deathist viewpoint tends to go with a certain element of hypocrisy. It is to be hoped and expected that a good many of death’s apologists, if they were one day presented with the concrete choice between (A) getting sick, old, and dying, and (B) being given a new shot of life to stay healthy, vigorous and to remain in the company of friends and loved ones to participate in the unfolding of the future, would, when push came to shove, choose this latter alternative.

If some people would still choose death, that’s a choice that is of course to be regretted, but nevertheless this choice must be respected. The transhumanist position on the ethics of death is crystal clear: death should be voluntary. This means that everybody should be free to extend their lives and to arrange for cryonic suspension of their deanimated bodies. It also means that voluntary euthanasia, under conditions of informed consent, is a basic human right.

It may turn out to be impossible to live forever, strictly speaking, even for those who are lucky enough to survive to such a time when technology has been perfected, and even under ideal conditions. The amount of matter and energy that our civilization can lay its hands on before they recede forever beyond our reach (due to the universe’s expansion) is finite in the current most favored cosmological models. The heat death of the universe is thus a matter of some personal concern to optimistic transhumanists!

It is too early to tell whether our days are necessarily numbered. Cosmology and fundamental physics are still incomplete and in theoretical flux; theoretical possibilities for infinite information processing (which might enable an upload to live an infinite life) seem to open and close every few years. We have to live with this uncertainty, along with the much greater uncertainty about whether any of us will manage to avoid dying prematurely, before technology has become mature.

4.5 Are transhumanist technologies environmentally sound?

The environmental impact of a technology depends on how it is used. Safeguarding the natural environment requires political will as well as good technology. The technologies necessary for realizing the transhumanist vision can be environmentally sound. Information technology and medical procedures, for example, tend to be relatively clean.

Transhumanists can in fact make a stronger claim regarding the environment: that current technologies are unsustainable. We are using up essential resources, such as oil, metal ores, and atmospheric pollution capacity, faster than they regenerate. At the present rate of consumption, we look set to exhaust these resources some time in this century. Any realistic alternatives that have been proposed involve taking technology to a more advanced level. Not only are transhumanist technologies ecologically sound, they may be the only environmentally viable option for the long term.

With mature molecular manufacturing [see “What is molecular nanotechnology?”], we will have a way of producing most any commodity without waste or pollution. Nanotechnology would also eventually make it economically feasible to build space-based solar plants, to mine extraterrestrial bodies for ore and minerals, and to move heavy industries off-earth. The only truly long-term solution to resource shortage is space colonization.

From a transhumanist point of view, humans and our artifacts and enterprises are part of the extended biosphere. There is no fundamental dichotomy between humanity and the rest of the world. One could say that nature has, in humanity, become conscious and self-reflective. We have the power to dream of a better ways for things to be and to deliberately set out to build our dreams, but we also have the responsibility to use this power in ways that are sustainable and that protect essential values.

5 TRANSHUMANISM AS A PHILOSOPHICAL AND CULTURAL VIEWPOINT

5.1 What are the philosophical and cultural antecedents of transhumanism?

The human desire to acquire posthuman attributes is as ancient as the human species itself. Humans have always sought to expand the boundaries of their existence, be it ecologically, geographically, or mentally. There is a tendency in at least some individuals always to try to find a way around every limitation and obstacle.

Ceremonial burial and preserved fragments of religious writings show that prehistoric humans were deeply disturbed by the death of their loved ones and sought to reduce the cognitive dissonance by postulating an afterlife. Yet, despite the idea of an afterlife, people still endeavored to extend life. In the Sumerian *Epic of Gilgamesh* (approx. 2000 B.C.), a king embarks on a quest to find an herb that can make him immortal. It's worth noting that it was assumed both that mortality was not inescapable in principle, and that there existed (at least mythological) means of overcoming it. That people really strove to live longer and richer lives can also be seen in the development of systems of magic and alchemy; lacking scientific means of producing an elixir of life, one resorted to magical means. This strategy was adopted, for example, by the various schools of esoteric Taoism in China, which sought physical immortality and control over or harmony with the forces of nature.

The Greeks were ambivalent about humans transgressing our natural confines. On the one hand, they were fascinated by the idea. We see it in the myth of Prometheus, who stole the fire from Zeus and gave it to the humans, thereby permanently improving the human condition. And in the myth of Daedalus, the gods are repeatedly challenged, quite successfully, by a clever engineer and artist, who uses non-magical means to extend human capabilities. On the other hand, there is also the concept of *hubris*: that some ambitions are off-limit and would backfire if pursued. In the end, Daedalus' enterprise ends in disaster (not, however, because it was punished by the gods but owing entirely to natural causes).

Greek philosophers made the first, stumbling attempts to create systems of thought that were based not purely on faith but on logical reasoning. Socrates and the sophists extended the application of critical thinking from metaphysics and cosmology to include the study of ethics and questions about human society and human psychology. Out of this inquiry arose cultural humanism, a very important current throughout the history of Western science, political theory, ethics, and law.

In the Renaissance, human thinking was awoken from medieval otherworldliness and the scholastic modes of reasoning that had predominated for a millennium, and the human being and the natural world again became legitimate objects of study. Renaissance humanism encouraged people to rely on their own observations and their own judgment rather than to defer in every matter to religious authorities. Renaissance humanism also created the ideal of the well-rounded personality, one that is highly developed scientifically, morally, culturally, and spiritually. A milestone is Giovanni Pico della Mirandola's *Oration on the Dignity of Man* (1486), which states that man does not have a ready form but that it is man's task to form himself. And crucially, modern science began to take form then, through the works of Copernicus, Kepler, and Galileo.

The Age of Enlightenment can be said to have started with the publication of Francis Bacon's *Novum Organum*, "the new tool" (1620), in which he proposes a scientific methodology based on empirical investigation rather than a priori reasoning. Bacon advocates the project of "effecting all things possible," by which he meant the achievement of mastery over nature in order to improve the condition of human beings.

The heritage from the Renaissance combines with the influences of Isaac Newton, Thomas Hobbes, John Locke, Immanuel Kant, Marquis de Condorcet, and others to form the basis for rational humanism, which emphasizes science and critical reasoning – rather than revelation and religious authority – as ways of learning about the natural world and the destiny and nature of man and of providing a grounding for morality. Transhumanism traces its roots to this rational humanism.

In the 18th and 19th centuries we begin to see glimpses of the idea that even humans themselves can be developed through the appliance of science. Benjamin Franklin and Voltaire speculated about extending human life span through medical science. Especially after Darwin's theory of evolution, atheism or agnosticism came to be seen as increasingly attractive alternatives. However, the optimism of the late 19th century often degenerated into narrow-minded positivism and the belief that progress was automatic. When this view collided with reality, some people reacted by turning to irrationalism, concluding that since reason was not sufficient, it was worthless. This resulted in the anti-technological, anti-intellectual sentiments whose sequelae we can still witness today in some postmodernist writers, in the New Age movement, and among the neo-Luddite wing of the anti-globalization agitators.

A significant stimulus in the formation of transhumanism was the essay *Daedalus: Science and the Future* (1923) by the British biochemist J. B. S. Haldane, in which he discusses how scientific and technological findings may come to affect society and improve the human condition. This essay set off a chain reaction of future-oriented discussions, including *The World, the Flesh and the Devil* by J. D. Bernal (1929), which speculates about space colonization and bionic implants as well as mental improvements through advanced social science and psychology; the works of Olaf Stapledon; and the essay "Icarus: the Future of Science" (1924) by Bertrand Russell, who took a more pessimistic view, arguing that without more kindness in the world, technological power will mainly serve to increase men's ability to inflict harm on one another. Science fiction authors such as H. G. Wells and Olaf Stapledon also got many people thinking about the future evolution of the human race. One frequently cited work is Aldous Huxley's *Brave New World* (1932), a dystopia where psychological conditioning, promiscuous sexuality, biotechnology, and opiate drugs are used to keep the population placid and contented in a static, totalitarian society ruled by an elite consisting of ten "world controllers". Huxley's novel warns of the dehumanizing potential of technology being used to arrest growth and to diminish the scope of human nature rather than enhance it.

The Second World War changed the direction of some of those currents that result in today's transhumanism. The eugenics movement, which had previously found advocates not only among racists on the extreme right but also among socialists and progressivist social democrats, was thoroughly discredited. The goal of creating a new and better world through a centrally imposed vision became taboo and passé; and the horrors of the Stalinist Soviet Union again underscored the dangers of such an approach. Mindful of these historical lessons, transhumanists are often deeply suspicious of collectively orchestrated change, arguing instead for the right of individuals to redesign themselves and their own descendants.

In the postwar era, optimistic futurists tended to direct their attention more toward technological progress, such as space travel, medicine, and computers. Science began to catch up with speculation. Transhumanist ideas during this period were discussed and analyzed chiefly in the literary genre of science fiction. Authors such as Arthur C. Clarke, Isaac Asimov, Robert Heinlein, Stanislaw Lem, and later Bruce Sterling, Greg Egan, and Vernor Vinge have explored various aspects of transhumanism in their writings and contributed to its proliferation.

Robert Ettinger played an important role in giving transhumanism its modern form. The publication of his book *The Prospect of Immortality* in 1964 led to the creation of the cryonics movement. Ettinger argued that since medical technology seems to be constantly progressing, and since chemical activity comes to a complete halt at low temperatures, it should be possible to freeze a person today and preserve the body until such a time when technology is advanced enough to repair the freezing damage and reverse the original cause of deanimation. In a later work, *Man into Superman* (1972), he discussed a number of conceivable improvements to the human being, continuing the tradition started by Haldane and Bernal.

Another influential early transhumanist was F. M. Esfandiary, who later changed his name to FM-2030. One of the first professors of future studies, FM taught at the New School for Social Research in New York in the 1960s and formed a school of optimistic futurists known as the UpWingers. In his book *Are you a transhuman?* (1989), he described what he saw as the signs of the emergence of the transhuman person, in his terminology indicating an evolutionary link towards posthumanity. (A terminological aside: an early use of the word “transhuman” was in the 1972-book of Ettinger, who doesn’t now remember where he first encountered the term. The word “transhumanism” was coined by Julian Huxley in *New Bottles for New Wine* (1957); the sense in which he used it, however, was not quite the contemporary one.)

In the 1970s and 1980s, several organizations sprung up for life extension, cryonics, space colonization, science fiction, and futurism. They were often isolated from one another, and while they shared similar views and values, they did not yet amount to any unified coherent worldview. One prominent voice from a standpoint with strong transhumanist elements during this era came from Marvin Minsky, an eminent artificial intelligence researcher.

In 1986, Eric Drexler published *Engines of Creation*, the first book-length exposition of molecular manufacturing. (The possibility of nanotechnology had been anticipated by Nobel Laureate physicist Richard Feynman in a now-famous after-dinner address in 1959 entitled “There is Plenty of Room at the Bottom”.) In this groundbreaking work, Drexler not only argued for the feasibility of assembler-based nanotechnology but also explored its consequences and began charting the strategic challenges posed by its development. Drexler’s later writings supplied more technical analyses that confirmed his initial conclusions. To prepare the world for nanotechnology and work towards it safe

implementation, he founded the Foresight Institute together with his then wife Christine Peterson in 1986.

Ed Regis's *Great Mambo Chicken and the Transhuman Condition* (1990) took a humorous look at transhumanism's hubristic scientists and philosophers. Another couple of influential books were roboticist Hans Moravec's seminal *Mind Children* (1988) about the future development of machine intelligence, and more recently Ray Kurzweil's bestselling *Age of Spiritual Machines* (1999), which presented ideas similar to Moravec's. Frank Tipler's *Physics of Immortality* (1994), inspired by the writings of Pierre Teilhard de Chardin (a paleontologist and Jesuit theologian who saw an evolutionary telos in the development of an encompassing noosphere, a global consciousness) argued that advanced civilizations might come to have a shaping influence on the future evolution of the cosmos, although some were put off by Tipler's attempt to blend science with religion. Many science advocates, such as Carl Sagan, Richard Dawkins, Steven Pinker, and Douglas Hofstadter, have also helped pave the way for public understanding of transhumanist ideas.

In 1988, the first issue of the *Extropy Magazine* was published by Max More and Tom Morrow, and in 1992 they founded the Extropy Institute (the term "extropy" being coined as an informal opposite of "entropy"). The magazine and the institute served as catalysts, bringing together disparate groups of people with futuristic ideas. More wrote the first definition of transhumanism in its modern sense, and created his own distinctive brand of transhumanism, "extropianism", which emphasized individualism, dynamic optimism, and the market mechanism in addition to technology. The transhumanist arts genre became more self-aware through the works of the artist Natasha Vita-More. During this time, an intense exploration of ideas also took place on various Internet mailing lists. Influential early contributors included Anders Sandberg (then a neuroscience doctoral student) and Robin Hanson (an economist and polymath) among many others.

The World Transhumanist Association was founded in 1998 by Nick Bostrom (who is its current Chair) and David Pearce to act as a coordinating international nonprofit organization for all transhumanist-related groups and interests, across the political spectrum. The WTA focuses on supporting transhumanism as a serious academic discipline and on promoting public awareness of transhumanist thinking. The WTA began publishing the *Journal of Evolution and Technology*, the first scholarly peer-reviewed journal for transhumanist studies in 1999 (which is also the year when the first version of this FAQ was published). In 2001, the WTA adopted its current constitution and is now governed by an executive board that is democratically elected by its full membership. James Hughes especially (the current WTA Secretary) among others helped lift the WTA to its current more mature stage, and a strong team of volunteers has been building up the organization to what it is today.

In the past couple of years, the transhumanist movement has been growing fast and furiously. Local groups are mushrooming in all parts of the world. Awareness of transhumanist ideas is spreading. Transhumanism is undergoing the transition from being the preoccupation of a fringe group of intellectual pioneers to becoming a mainstream

approach to understanding the prospects for technological transformation of the human condition. That technological advances will help us overcome many of our current human limitations is no longer an insight confined to a few handfuls of techno-savvy visionaries. Yet understanding the consequences of these anticipated possibilities and the ethical choices we will face is a momentous challenge that humanity will be grappling with over the coming decades. The transhumanist tradition has produced a (still evolving) body of thinking to illuminate these complex issues that is unparalleled in its scope and depth of foresight.

References:

Bacon, F. *Novum Organum*. (New York: Colonial Press, 1899 [1620]).
http://www.constitution.org/bacon/nov_org.htm

Bernal, J. D. *The World, the Flesh & the Devil: An Enquiry into the Future of the Three Enemies of the Rational Soul*. (Bloomington: Indiana University Press, 1969 [1929]).
<http://www.santafe.edu/~shalizi/Bernal/>

Drexler, E. *The Engines of Creation: The Coming Era of Nanotechnology*. (New York: Anchor Books, 1986). <http://www.foresight.org/EOC/index.html>

Ettinger, R. *The Prospect of Immortality*. (New York: Doubleday, 1964).
<http://www.cryonics.org/book1.html>

Ettinger, R. *Man into Superman*. (New York: St. Martin's, 1972). <http://www.cryonics.org/book2.html>

Extropy Institute. <http://www.extropy.org>

Feynman, R. "There is Plenty of Room at the Bottom." Presentation given on December 29th, 1959 at the annual meeting of the American Physical Society at the California Institute of Technology, published in *Engineering and Science*, Feb 1960. <http://www.zyvex.com/nanotech/feynman.html>

FM-2030. *Are You a Transhuman?* (New York: Warner Books, 1989).

Foresight Institute. <http://www.foresight.org>

Haldane, J. B. S. *Daedalus or Science and the Future*. (New York: E. P. Dutton & Co., Inc., 1924 [1923]).
<http://www.santafe.edu/~shalizi/Daedalus.html>

Huxley, A. *Brave New World*. (San Bernadino: The Borgo Press, 1989 [1932]).

Huxley, J. *New Bottles for New Wine*. (New York: Harper, 1957).

Journal of Evolution and Technology. <http://www.jetpress.org/>

Mirandola, Giovanni Pico. *Oration on the Dignity of Man*. (1486).
<http://www.santafe.edu/~shalizi/Mirandola/>

Moravec, H. *Mind Children* (Harvard: Harvard University Press, 1988).

Regis, E. *Great Mambo Chicken and the Transhuman Condition* (New York: Perseus, 1990).

Russell, B. *Icarus or The Future of Science*. (New York: E. P Dutton & Company, 1924).
<http://www.santafe.edu/~shalizi/Icarus.html>

Tipler, F. *The Physics of Immortality* (New York: Doubleday, 1994).

World Transhumanist Association. <http://www.transhumanism.org>

5.2 What currents are there within transhumanism? Is extropianism the same as transhumanism?

There is a rich variety of opinion within transhumanist thought. Many of the leading transhumanist thinkers hold complex and subtle views that are under constant revision and development and which often defy easy labeling. Some distinctive – although not always sharply defined – currents or flavors of transhumanism can nevertheless be discerned:

Extropianism. The name is derived from the term “extropy”, coined by T. O. Morrow in 1988, referring to “the extent of a system's intelligence, information, order, vitality, and capacity for improvement”. Extropianism is defined by the Extropian Principles, a text authored by Max More (1998), who co-founded the Extropy Institute together with Morrow. Version 3.0 of this document lists seven principles that are important for extropians in the development of their thinking: Perpetual Progress, Self-Transformation, Practical Optimism, Intelligent Technology, Open Society, Self-Direction, and Rational Thinking. These are meant to codify general attitudes rather than specific dogmas.

Democratic transhumanism. This strand of transhumanism advocates both the right to use technology to transcend the limitations of the human body and the extension of democratic concerns beyond formal legal equality and liberty, into economic and cultural liberty and equality, in order to protect values such as equality, solidarity, and democratic participation in a transhuman context (Hughes 2002).

The Hedonistic Imperative. Another transhumanist current is represented by advocates of “paradise-engineering” as outlined in David Pearce (2003). Pearce argues on ethical grounds for a biological program to eliminate all forms of cruelty, suffering, and malaise. In the short-run, our emotional lives might be enriched by designer mood-drugs (i.e. not street-drugs). In the long-term, however, Pearce suggests that it will be technically feasible to rewrite the vertebrate genome, redesign the global ecosystem, and use biotechnology to abolish suffering throughout the living world. Pearce believes “post-Darwinian superminds” will enjoy genetically pre-programmed well-being and be animated by “gradients of bliss”.

Singularitarianism. Singularitarian transhumanists focus on transhuman technologies that can potentially lead to the rise of smarter-than-human intelligence, such as brain-computer interfacing and Artificial Intelligence. Since our present-day intelligence is ultimately the source of our technology, singularitarians expect the technological creation of smarter-than-human intelligence to be a watershed moment in history, with an impact more comparable to the rise of Homo sapiens than to past breakthroughs in technology. Singularitarians stress the importance of ensuring that such intelligence be coupled with ethical sensibility (Yudkowsky 2003) [see also “What is the singularity?”].

Theoretical transhumanism. This is not so much a specific version of a transhumanism as a research direction: the study of the constraints, possibilities, and consequences of potential future trajectories of technological and human development, using theoretical tools from economics, game theory, evolution theory, probability theory, and “theoretical applied science” i.e. the study of physically possible systems designs that we cannot yet build. For some examples, see Bostrom (2002, 2003a) and Hanson (1994, 1998). Investigations of ethical issues related to the transhumanist project – the project of creating a world where as many people as possible have the option of becoming posthuman – can also be included under this heading (see e.g. Bostrom 2003b).

Salon transhumanism. Transhumanism as a network of people who share certain interests and like to spend long hours conversing about transhumanist matters on email lists or face-to-face.

Transhumanism in arts and culture. Transhumanism as a source of inspiration in artistic creation and cultural activities, including efforts to communicate transhumanist ideas and values to a wider audience [see also “What kind of transhumanist art is there?”].

References:

- Bostrom, N. “Existential Risks: Analyzing Human Extinction Scenarios.” *Journal of Evolution and Technology*. (2002), Vol. 9. <http://jetpress.org/volume9/risks.html>
- Bostrom, N. “Are You Living In A Computer Simulation?” *Philosophical Quarterly*. (2003a), Vol. 53, No. 211, pp. 243-255. <http://www.simulation-argument.com/simulation.html>
- Bostrom, N. “Human Genetic Enhancements: A Transhumanist Perspective.” *The Journal of Value Inquiry*. (2003b), forthcoming.
- Hanson, R. “What if Uploads Come First: The Crack of a Future Dawn.” *Extropy*, Vol. 6, No. 2 (1994). <http://hanson.gmu.edu/uploads.html>
- Hanson, R. “Burning the Cosmic Commons: Evolutionary Strategies for Interstellar Colonization.” (1998). <http://hanson.gmu.edu/filluniv.pdf>
- Hughes, J. “Democratic Transhumanism.” *Transhumanity*, April 28, 2002. http://www.transhumanism.com/articles_more.php?id=P52_0_4_0_C
- Pearce, D. *The Hedonistic Imperative* (version of 2003). <http://www.hedweb.com/hedethic/hedonist.htm>
- More, M. “The Extropian Principles, v. 3.0.” (1998). <http://www.maxmore.com/extprn3.htm>
- Yudkowsky, E. “What is the Singularity.” (2003). <http://www.singinst.org/what-singularity.html>

5.3 How does transhumanism relate to religion?

Transhumanism is a philosophical and cultural movement concerned with promoting responsible ways of using technology to enhance human capacities and to increase the scope of human flourishing.

While not a religion, transhumanism might serve a few of the same functions that people have traditionally sought in religion. It offers a sense of direction and purpose and suggests a vision that humans can achieve something greater than our present condition. Unlike most religious believers, however, transhumanists seek to make their dreams come true in *this* world, by relying not on supernatural powers or divine intervention but on rational thinking and empiricism, through continued scientific, technological, economic, and human development. Some of the prospects that used to be the exclusive thunder of the religious institutions, such as very long lifespan, unfading bliss, and godlike intelligence, are being discussed by transhumanists as hypothetical future engineering achievements.

Transhumanism is a naturalistic outlook. At the moment, there is no hard evidence for supernatural forces or irreducible spiritual phenomena, and transhumanists prefer to derive their understanding of the world from rational modes of inquiry, especially the scientific method. Although science forms the basis for much of the transhumanist worldview, transhumanists recognize that science has its own fallibilities and imperfections, and that critical ethical thinking is essential for guiding our conduct and for selecting worthwhile aims to work towards.

Religious fanaticism, superstition, and intolerance are not acceptable among transhumanists. In many cases, these weaknesses can be overcome through a scientific and humanistic education, training in critical thinking, and interaction with people from different cultures. Certain other forms of religiosity, however, may well be compatible with transhumanism.

It should be emphasized that transhumanism is not a fixed set of dogmas. It is an evolving worldview, or rather, a family of evolving worldviews – for transhumanists disagree with each other on many issues. The transhumanist philosophy, still in its formative stages, is meant to keep developing in the light of new experiences and new challenges. Transhumanists want to find out where they are wrong and to change their views accordingly.

5.4 Won't things like uploading, cryonics, and AI fail because they can't preserve or create the soul?

If we answer this question from a religious standpoint, there is no clear ground for ruling out these technologies as incompatible with teachings about the soul. There is no scriptural basis in the Bible for assuming that God can't get to our soul if we freeze our physical body, nor is there a single word in the Christian or Jewish scriptures, or the Quran, the Dhammapada, or the Tao Teh Ching, that prohibits cryonics. Or, for someone who believes in reincarnation, there are no traditional beliefs that say reincarnation is prevented when someone freezes to death or whose body is frozen after clinical death. If there is a soul and it enters the body at conception, then cryonics may well work – after all, human embryos have been frozen, stored for extended periods, and then implanted in

their mothers, resulting in healthy children (who presumably have souls). Uploading and machine intelligence may reveal new things to us about the soul works. It is interesting to note that the Dalai Lama, when asked, did not rule out the possibility of reincarnating into computers (Hayward et al. 1992), pp. 152f.

While the concept of a soul is not used much in a naturalistic philosophy such as transhumanism, many transhumanists do take an interest in the related problems concerning personal identity (Parfit 1984) and consciousness (Churchland 1988). These problems are being intensely studied by contemporary analytic philosophers, and although some progress has been made, e.g. in Derek Parfit's work on personal identity, they have still not been resolved to general satisfaction.

References:

Churchland, P. *Matter and Consciousness*. (Cambridge, MA: MIT Press, 1988).

Hayward, J. et al. *Gentle Bridges: Conversations with the Dalai Lama on the Sciences of the Mind*. (Shambala Publications, 1992).

Parfit, D. *Reasons and Persons*. (Oxford: Oxford University Press, 1984).

5.5 What kind of transhumanist art is there?

Many kinds, but what examples one would give depends on how one defines "transhumanist art". If one defines it simply as art that is concerned with the human aspiration to overcome current limits, then a large portion of all art through the ages would count as transhumanist – from ancient myths of Promethean hubris, to religious transcendental iconography, architecture, and rituals, J. S. Bach's fugues, Goethe's Faust, through to the postmodern artists, many of whom conceived of their work as an attempt to explode conceptual barriers in order to widen the reach of human creativity.

If we narrow the definition by adding the requirement that a transhumanist telos be coupled to a notion of the centrality of technological means, we get a different set of paradigmatic examples. The Frankenstein myth (based originally on the novel by Mary Shelley published in 1831, and elaborated in countless forms since then) is one classic, and in general science fiction has been the genre most intensely preoccupied with transhumanist themes, reaching back to Jules Verne and Karel Čapek, through Isaac Asimov, Robert A. Heinlein, Stanislav Lem, Arthur C. Clark, on to Vernor Vinge, Bruce Sterling, James Halperin, Greg Egan, and many others. Many of these author's stories have been adapted for the screen. (The Star Trek series features cool new technology but the same old humans, so it is not a very paradigmatic exemplar of transhumanist art.)

Another way of defining the concept would be to say that transhumanist art is art produced by declared transhumanists. On this definition, examples have to be sought in recent times since the term "transhumanism" in its contemporary sense is quite new. Natasha Vita-More is one of the earliest and most prominent transhumanist artists in this sense. For instance, her recent visual and conceptual work, Primo 3M+, presents a kind of sleek future shopping catalog entry for an entire body design with features such as

memory enhancements, sonar sensors, solar protected skin with hue-texture changeability, gender reconfigurability, environmentally-friendly waste disposal, and which comes complete with warranty and upgradability. Vita-More is also the author of several transhumanist arts manifestos, in which transhumanist art becomes self-conscious for the first time. Other contemporary transhumanist artists include Lilia Morales y Mori, Anders Sandberg, Juan Meridalva (visual art); Elaine Walker, E. Shaun Russell, Emlyn O'Regan, Gustavo Muccillo Alves, and the band Cosmodelia (electronic music); Susan Rogers (puppet theatre); Jane Holt (performance art); and many others.

References:

Vita-More, N. *Primo 3M+* (2002). <http://www.natasha.cc/primo.htm>

Vita-More, N. "Transhumanist Arts Statement" (version of 2002). <http://www.extropic-art.com/transart.htm>

6 PRACTICALITIES

6.1 What are the reasons to expect all these changes?

Take a look around. Compare what you see with what you would have seen only fifty years ago. It is not an especially bold conjecture that the next 50 years will see at least as much change and that the state of technology in the mid-21st century will be quite wondrous by present standards. The conservative projection, which assumes only that progress continues in the same gradual way it has since the 17th century, would imply that we should expect to see dramatic developments over the coming decades.

This expectation is reinforced when one considers that many crucial areas seem poised for critical breakthroughs. The World-Wide Web is beginning to link the world's people, adding a new global layer to human society where information is supreme. The Human Genome Project has been completed, and the study of the functional roles of our genes (functional genomics) is proceeding rapidly. Techniques for using this genetic information to modify adult organisms or the germ-line are being developed. The performance of computers doubles every 18 months and will approach the computational power of a human brain in the foreseeable future. Pharmaceutical companies are refining drugs that will enable us to regulate mood and aspects of personality with few side effects. Many transhumanist aims can be pursued with present technologies. Can there be much doubt that, barring a civilization-destroying cataclysm, technological progress will give us much more radical options in the future? [See also "Won't these developments take thousands or millions of years?"]

Molecular manufacturing has the potential to transform the human condition. Is it a feasible technology? Eric Drexler and others have showed in detail how machine-phase nanotechnology is consistent with physical laws and have outlined several routes by

which it could be developed [see “What is molecular nanotechnology?”]. Molecular manufacturing might seem incredible, maybe because the eventual consequences seem too overwhelming, but nanotechnology experts point out that there currently exists no published technical critique of Drexler’s arguments. More than ten years after the publication of *Nanosystems*, nobody has yet been able to point to any significant error in the calculations. Meanwhile, investment in the development of nanotechnology, already billions of dollars annually worldwide, is growing every year, and at least the less visionary aspects of nanotechnology have already become mainstream.

There are many independent methods and technologies that can enable humans to become posthuman. There is uncertainty about which technologies will be perfected first, and we have a choice about which methods to use. But provided civilization continues to prosper, it seems almost inevitable that humans will sooner or later have the option of becoming posthuman persons. And, unless forcibly prevented, many will choose to explore that option.

References:

Drexler, E. *Nanosystems: Molecular Machinery, Manufacturing, and Computation*. (New York: John Wiley & Sons, 1992).

6.2 Won’t these developments take thousands or millions of years?

It is often very hard to predict how long a certain technological development will take. The moon landing happened sooner than most people had expected, but fusion energy still eludes us after half a century of anticipation. The difficulty in forecasting the timing lies partly in the possibility of unexpected technical obstacles and partly in the fact that the rate of progress depends on levels of funding, which in turn depends on hard-to-predict economic and political factors. Therefore, while one can in many cases give good grounds for thinking that a technology will eventually be developed, one can usually only make informed guesses about how long it will take.

The vast majority of transhumanists think that superintelligence and nanotechnology will both be developed in less than a hundred years, and many predict that it will happen well within the first third of this century. (Some of the reasons for holding these opinions are outlined in the sections about these two technologies.) Once there is both nanotechnology and superintelligence, a very wide range of special applications will follow swiftly.

It would be possible to give a long list of examples where people in the past have solemnly declared that something was technologically absolutely impossible,

“The secrets of flight will not be mastered within our lifetime – not within a thousand years.” (Wilbur Wright, 1901),

or socially irrelevant,

“There is no reason why anyone would want a computer in their home.” (Ken Olsen: President, Chairman and Founder of Digital Equipment Corporation, 1977)

– only to see it happen few years later. However, one could give an equally long list of cases of predicted breakthroughs that failed to occur. The question cannot be settled by enumerating historical parallels.

A better strategy is to look directly at what a careful analysis of the underlying physical constraints and engineering constraints might reveal. In the case of the most crucial future technologies – superintelligence and molecular manufacturing – such analyses have been done. Many experts believe that these will likely be achieved within the first several decades of the 21st century. Other experts think it will take much longer. There seems to be more disagreement about the feasibility and time-frame of superintelligence than of nanotechnology.

Another way of forming a view of where we are headed is by looking at trends. At least since the late 19th century, science and technology, as measured by a wide range of indicators, have doubled about every 15 years (Price 1986). Extrapolating this exponential rate of progress, one is led to expect to see dramatic changes in the relatively near future. It would require an abrupt reversal of current trends, an unexpected deceleration, in order for the changes that many transhumanists foresee not to happen within the 21st century.

References:

The Foresight Institute. “Erroneous Predictions and Negative Comments Concerning Scientific and Technological Developments.” (2002). <http://www.foresight.org/News/negativeComments.html>

Price, D. J. *Little Science, Big Science ...and Beyond*. (New York: Columbia University Press, 1986).

6.3 What if it doesn’t work?

Success in the transhumanist endeavor is not an all-or-nothing matter. There is no “it” that everything hinges on. Instead, there are many incremental processes at play, which may work better or worse, faster or more slowly. Even if we can’t cure all diseases, we will cure many. Even if we don’t get immortality, we can have healthier lives. Even if we can’t freeze whole bodies and revive them, we can learn how to store organs for transplantation. Even if we don’t solve world hunger, we can feed a lot of people. With many potentially transforming technologies already available and others in the pipeline, it is clear that there will be a large scope for human augmentation. The more powerful transhuman technologies, such as machine-phase nanotechnology and superintelligence, can be reached through several independent paths. Should we find one path to be blocked, we can try another one. The multiplicity of routes adds to the probability that our journey will not come to a premature halt.

There *are* ways to fail completely, namely if we succumb to an existential disaster [see “Aren’t these future technologies very risky? Could they even cause our extinction?”]. Efforts to reduce existential risks are therefore a top priority.

6.4 How can I use transhumanism in my own life?

While transhumanism has been known to cross over with academic agendas, ethical philosophies, political causes, and artistic movements, transhumanism is not a lifestyle, a religion, or a self-help guide. Transhumanism can’t tell you what kind of music to listen to, which hobbies to pursue, whom to marry or how to live your life, any more than, say, being a member of Amnesty International or studying molecular biology could tell you these things.

Depending on your situation and your needs, you might or might not find some of the currently available human modification or enhancement options useful. Some of these are commonplace – exercise, healthy diet, relaxation techniques, time management, study skills, information technology, coffee or tea (as stimulants), education, and nutritional supplements (such as vitamins, minerals, fatty acids, or hormones). Others you might not have thought of, such as getting a cryonic suspension contract [see “What is cryonics? Isn’t the probability of success too small?”], or chewing nicotine gum for its nootropic effects. Still others – for instance pharmacological mood drugs or sex reassignment surgery – are suitable only for people who have special difficulties or needs.

If you want to learn more about transhumanist topics, meet like-minded individuals, and participate in some way the transhumanist effort, see [“How can I get involved and contribute?”].

6.5 How could I become a posthuman?

At present, there is no manner by which any human can become a posthuman. This is the primary reason for the strong interest in life extension and cryonics among transhumanists. Those of us who live long enough to witness currently foreseeable technologies come to fruition may get the chance to become posthuman. Although there are no guarantees of success, there are some things that can be done on an individual level that will improve the odds a bit:

1. Live healthily and avoid unnecessary risks (such as smoking);
2. Sign up for cryonics;
3. Keep abreast of current research and save some money so that you can afford future life-extension treatments when they become available;
4. Support the development of transhuman technologies through donations, advocacy, investment, or choosing a career in the field; work to make access more universal and to make the world safer from existential risks [see “Aren’t these future technologies very risky? Could they even cause our extinction?”];

5. Join others to help promote transhumanism.

Meanwhile, we can enjoy and make the most of the opportunities that exist today for living worthwhile and meaningful lives. If we compare our current lot with that of our historical ancestors, most (at least those of us who don't live in the least developed countries) will find that the material circumstances for human flourishing are the best they have ever been. In addition, we possess an unprecedented accumulation of cultural and intellectual treasures whereby we can enrich our experiences and broaden our horizons.

6.6 Won't it be boring to live forever in a perfect world?

Why not try it and see?

"Perfection" is a vague and treacherous word. There is considerable disagreement among transhumanists about what kind of perfection is attainable and desirable, either in theory or in practice. It is probably wiser to speak of improving the world, rather than making it "perfect". Would it be boring to live for an indefinitely long time in a greatly improved world? The world could surely be improved over the way it is now, including becoming less boring. If you got rid of the pain and stress associated with, say, filling out annual tax returns, people would probably not sit around afterward saying: "Life feels meaningless now that I no longer have income tax forms to fill out."

Admittedly, material improvements to the environment may not, in themselves, be sufficient to bring about lasting happiness. If your accustomed fare is bread and water, then a box of cookies can be a feast. But if every night you eat out at fancy restaurants, such fine fare will soon seem ordinary and normal; and any lesser feast, such as a box of cookies, would be insulting by comparison. Some cognitive scientists speculate that we each have a "set point" of happiness, to which we soon return regardless of changes in the environment. There may be considerable truth to the folk wisdom that an expensive new car does not make you happier (or rather, it makes you happier, but only temporarily). In some ways, human minds and brains are just not designed to be happy. Fortunately, there are several potential viewpoints from which to go about addressing this challenge.

Apes engage in activities that we, as humans, would find repetitive and dull. In the course of becoming smarter, we have become bored by things that would have interested our ancestors. But at the same time we have opened up a vast new space of possibilities for having fun – and the new space is much larger than the previous one. Humans are not simply apes who can obtain more bananas using our intelligence as a tool. Our intelligence enables us to desire new things, such as art, science, and mathematics. If at any point in your indefinitely long life you become bored with the greatly improved world, it may only indicate that the time has come to bump up your intelligence another increment.

If the human brain has a “set point” of happiness to which it returns, maybe this is a design flaw and should be fixed – one of those things that we will end up defining as human, but not humane. It would probably be unwise to eliminate boredom entirely, since boredom can serve to prevent us from wasting too much time on monotonous and meaningless activities. But if we’re doing new things, learning, growing more intelligent, and we still aren’t happy, for no better reason than that our cognitive architecture is badly designed, then perhaps it is time to redesign it. Present clinical mood-drugs are crude, but nonetheless they can sometimes restore interest and enthusiasm for life – sometimes tiredness and despair has no interesting reason behind it and is simply an imbalance of brain chemistry. Only by compartmentalizing our thinking to a high degree can we imagine a world where there is mature molecular nanotechnology and superhuman artificial intelligence, but the means are still lacking to control the brain circuitry of boredom. Fundamentally, there is no reason why pleasure, excitement, profound well-being and simple joy at being alive could not become the natural, default state of mind for all who desire it.

Ed Regis (1990, p. 97) suggests the following points also be considered:

1. Ordinary life is sometimes boring. So what?
2. Eternal life will be as boring or as exciting as you make it.
3. Is being dead more exciting?
4. If eternal life becomes boring, you will have the option of ending it at any time.

Transhumanism is not about a fancier car, more money, or clever gadgetry, even though this is what the media presents to us as “science” and “advanced technology”; transhumanism is about genuine changes to the human condition, including increased intelligence and minds better suited to the achievement of happiness.

References:

Pearce, D. *The Hedonistic Imperative*. (2003) <http://www.hedweb.com>

Regis, E. *Great Mambo Chicken and the Transhuman Condition*. (Penguin Books: New York, 1990).

6.7 How can I get involved and contribute?

You can join the World Transhumanist Association. The WTA is a nonprofit, democratic membership organization that works to promote discussion of possibilities for the radical improvement of human capacities using technology, as well as of the ethical issues and risks involved in technological developments. It was founded in 1998 as an umbrella organization to publicize transhumanist ideas and to seek academic acceptance of transhumanism as a philosophical and cultural movement. The WTA organizes conferences, publishes an academic journal and a webzine, issues press statements, and coordinates student campus chapters and local transhumanist groups around the world.

To find out about current projects and upcoming events, and to become a member, please visit the WTA website.

The WTA has been growing since its inception and especially rapidly in the last couple of years, but the task before us is both momentous and mountainous. Your help is needed. There are myriad ways to contribute – organizing or participating in a local discussion group, writing articles or letters to the editor, making a financial contribution, spreading the word to friends and acquaintances, volunteering your skills to the WTA, translating key documents into other languages, linking to the WTA from your website, attending conferences and sharing your ideas, directing your research or creative activity towards transhumanist themes, to name but a few.

If you want to study transhumanist ideas in more detail, you can find some syllabi and reading lists on the WTA website to get you started. If you want to exchange ideas with others, or just listen in to ongoing conversations, you may want to join one of the mailing lists and newsgroups maintained by the WTA.

The coming technological transitions may be the most important challenge that humanity will ever face. The entire future of intelligent life on Earth may depend on how we handle it. If we do the right things, a wonderful posthuman future with limitless opportunities for growth and flourishing may lie ahead. If we handle it badly, intelligent life might go extinct. Don't you want to take part and attempt to make a difference for the better?

References:

World Transhumanist Association. <http://www.transhumanism.org>. (From this site, links to local groups and affiliated organizations can also be found.)

7 ACKNOWLEDGEMENTS AND DOCUMENT HISTORY

The Transhumanist FAQ was conceived as an attempt to develop a broadly based consensus articulation of the basics of responsible transhumanism. The aim was a text that could serve both as a guide to those new to the field and as a reference work for more seasoned participants.

Close to a hundred people have contributed in some way in the making of this document. The current version is a thoroughly revision of the version of 1999. Considerable new material has been added and many old sections have been substantially reworked. In preparing version 2.0, the following people have been especially helpful:

Eliezer Yudkowsky, who provided editorial assistance with the entire document and many comments on particular issues of substance; Dale Carrico who proofread the first

half of the text and eliminated a whole army of typos and infelicities; and Michael LaTorra who did the same for the second half; and “Reason” who then went over the whole document again, as did Frank Forman, and Sarah Banks Forman. Useful comments of either substance or form have also been contributed by (in alphabetical order): Michael Anissimov, Samantha Atkins, Milan Cirkovic, José Luis Cordeiro, George Dvorsky, James Hughes, G.E. Jordan, Vasso Kambourelli, Michael LaTorra, Eugen Leidl, Juan Meridalva, Harvey Newstrom, Emlyn O’Reagan, Christine Peterson, Giulio Prisco, Reason, Rafal Smigrodzki, Simon Smith, Mike Treder, and Mark Walker. Many others have over the years offered questions or reflections that have in some way helped shape this document, and even though it is not possible to name you all, your contributions are warmly appreciated. All remaining errors are my own fault.

Since this new version builds on its predecessor, the contributors to the latter are also indirect contributors to the present document. The people who assisted with the first version included especially Anders Sandberg (several of the original sections were based heavily on his material and the section on transhumanism’s historical precedents still very distinctly bears his mark), Kathryn Aegis, and Natasha Vita-More (who also both provided substantial chunks of text). The presentation in the cryonics section was, and still is, directly inspired by an article by Ralph Merkle. Several people contributed to the definition of transhumanism, in particular Kathryn Aegis and Max More. Greg Burch, David Pearce, Kathryn Aegis, and Anders Sandberg kindly offered extensive editorial comments. Ideas, criticisms, questions, phrases, and sentences to the original version were contributed by (in alphabetical order): Kathryn Aegis, Alex (intech@intsar.com), Brent Allsop, Brian Atkins, Scott Badger, Doug Bailey, Harmony Baldwin, Damien Broderick, Greg Burch, David Cary, John K Clark, Dan Clemmensen, Damon Davis, Jeff Dee, Jean-Michel Delhotel, Dylan Evans, EvMick@aol.com, Daniel Fabulich, Frank Forman, Robin Hanson, Andrew Hennessey, Tony Hollick, Joe Jenkins, William John, Michelle Jones, Arjen Kamphius, Henri Kluytmans, Eugene Leidl, Michael Lorrey, mark@unicorn.com, Peter C. McCluskey, Erik Moeller, J. R. Molloy, Max More, Bryan Moss, Harvey Newstrom, Michael Nielsen, John S. Novak III, Dalibor van den Otter, David Pearce, pilgrim@cyberdude.com, Thom Quinn, Anders Sandberg, Wesley R. Schwein, Shakeip@aol.com, Allen Smith, Geoff Smith, Randy Smith, Dennis Stevens, Derek Strong, Remi Sussan, Natasha Vita-More, Michael Wiik, Eliezer Yudkowsky, and zebo@pro-ns.net

I would like to thank you all for helping creating this FAQ and for making transhumanism possible!

Nick Bostrom
Oxford, October 2003

ON the bank at the end
Of what was there before us
Gazing over to the other side
On what we can become
Veiled in the mist of naïve speculation
We are busy here preparing
Rafts to carry us across

Before the light goes out leaving us
In the eternal night of could-have-beens